

SEARCH REQUEST FORM

Scientific and Technical Information Center

Access DB# _____

Requester's Full Name: DATRICIA DUFFY Examiner #: 72448 Date: 2/6/03
 Art Unit: 1645 Phone Number 305-7555 Serial Number: 09/833,799
 Mail Box and Bldg/Room Location: 8E12 Results Format Preferred (circle): PAPER DISK E-MAIL

If more than one search is submitted, please prioritize searches in order of need.

 Please provide a detailed statement of the search topic, and describe as specifically as possible the subject matter to be searched. Include the elected species or structures, keywords, synonyms, acronyms, and registry numbers, and combine with the concept or utility of the invention. Define any terms that may have a special meaning. Give examples or relevant citations, authors, etc, if known. Please attach a copy of the cover sheet, pertinent claims, and abstract.

Title of Invention: _____

Inventors (please provide full names): _____

Earliest Priority Filing Date: 6/8/90

For Sequence Searches Only Please include all pertinent information (parent, child, divisional, or issued patent numbers) along with the appropriate serial number.

09/833,799 has priority to 07/534,096 filed 6/8/90 and does not have to comply with the sequence rules.

Jan Delaval
 Reference Librarian
 Biotechnology & Chemical Library
 CM1 1E07-703-308-4498
 jan.delaval@uspto.gov

Please hand enter polypeptide and coding NA structure set forth in Fig 13 attached.

Please search ① NA + polypeptide and oligomers of each.
 ② AA run against NA data bases
 ③ interference search

Please print out top 75 hits in each category

Thanks

2/12/03 - 01 - 217 233 - 214 326
 63 - 30-49 - 11-31

Pat Duffy

3/16/03 - 01 - 243-703

2/12/03 - 01 - 458-478
 2/12/03 - 01 - 416-434
 06 - 248-304 - 7 8 Feb 03
 03 - 117-137

2/12/03 - 01 - 327-347
 03 - 80-100
 2/12/03 - 01 -
 02 - 64-84

STAFF USE ONLY

Searcher: 4495
 Searcher Phone #: _____
 Searcher Location: 211103
 Date Searcher Picked Up: 2/11/03
 Date Completed: _____
 Searcher Prep & Review Time: _____
 Clerical Prep Time: 45
 Online Time: 120

Type of Search

NA Sequence (#) ☒
 AA Sequence (#) _____
 Structure (#) _____
 Bibliographic _____
 Litigation _____
 Fulltext _____
 Patent Family _____
 Other _____

Vendors and cost where applicable

STN _____
 Dialog _____
 Questel/Orbit _____
 Dr. Link _____
 Lexis/Nexis _____
 Sequence Systems: ☒
 WWW/Internet _____
 Other (specify) _____

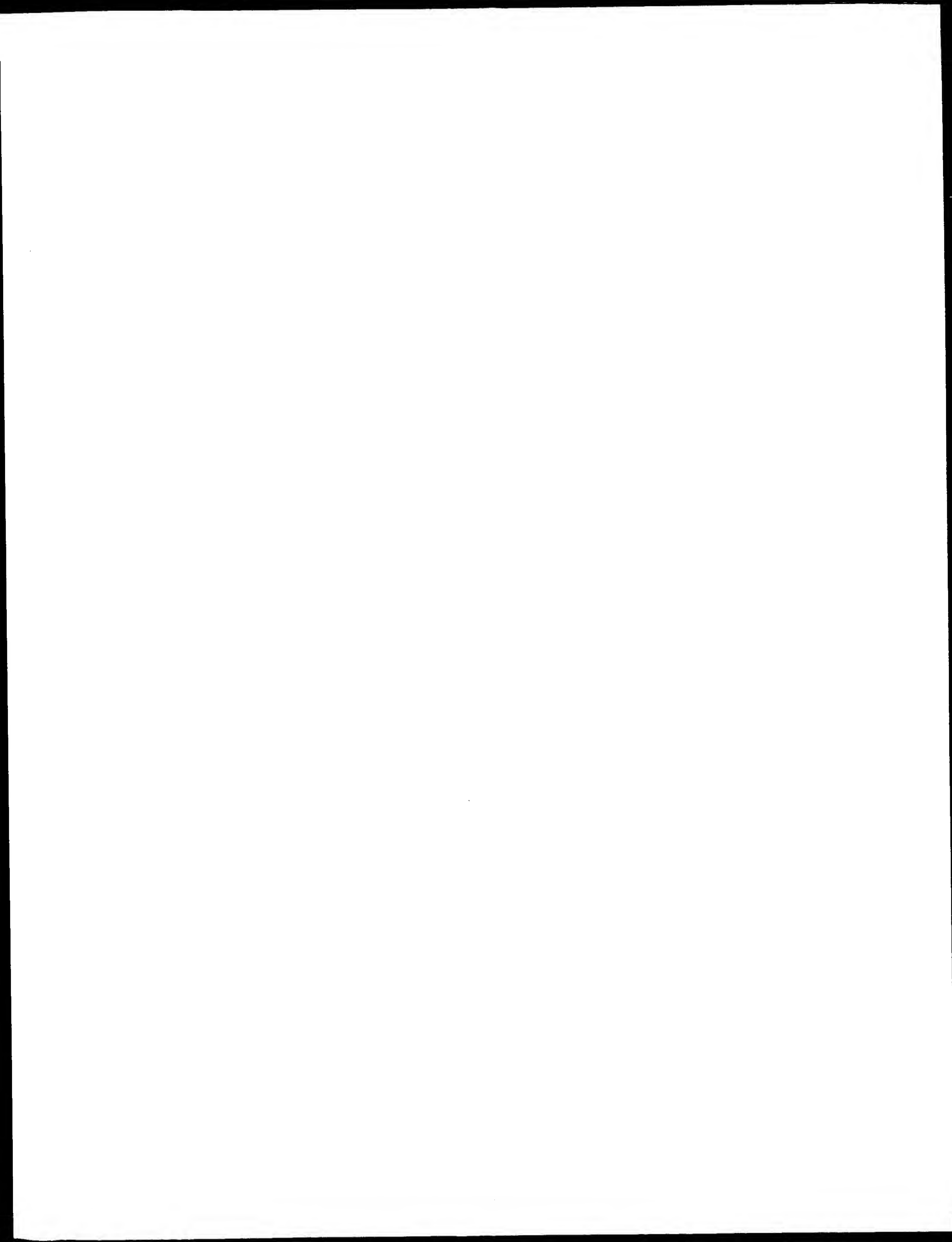


Fig. 13.

AlaGlnGluProValLysGlyProValSerThr

1 ELI1
AATTCGAGCTCGGTACCATACCTGCATATGCTCAAGAACCAGTTAAAGGTCCTGTGTCTACT
GCTCGAGCCATGGTATGGACGTATACGAGTTCTTGGTCAATTTCCAGGACACAGATGA

LysProGlySerCysProIleIleLeuIleArgCysAlaMetLeuAsnProProAsnArg

63 ELI3
AAGCCAGGTTCTTGTCTATTATCTTGATTCGTTGCGCTATGTTAAACCCACCTAACCGT
TTCGGTCCAAGAACAGGATAATAGAACTAAGCAACGCGATACAATTTGGGTGGATTGGCA
ELI2 ←

CysLeuLysAspThrAspCysProGlyIleLysLysCysCysGluGlySerCysGlyMet

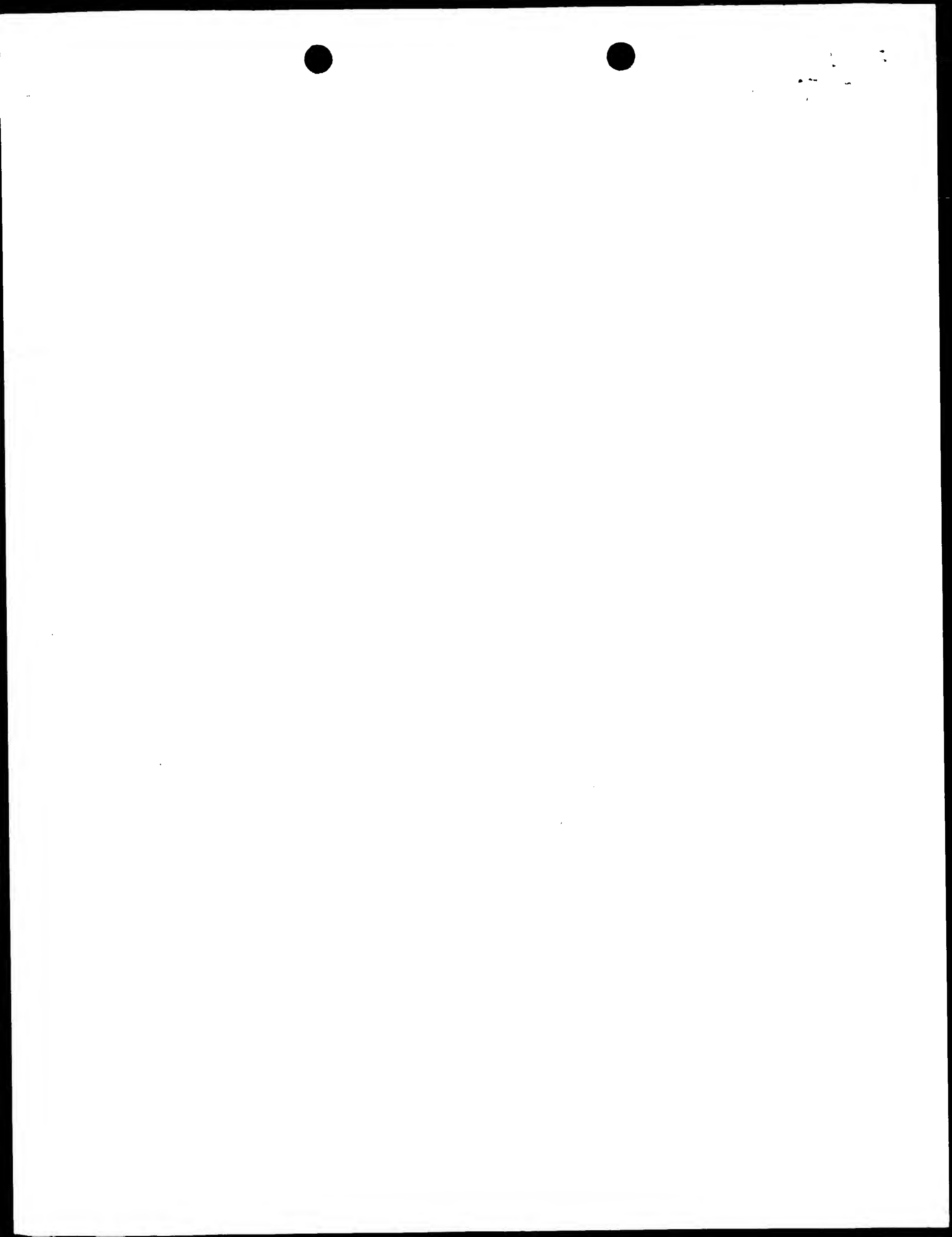
123 ELI5
TGTTTGAAGGACACTGATTGTCCAGGTATCAAAAAGTGCTGTGAAGGTTCTGCGGTATG
ACAACTTCCTGTGACTAACAGGTCCATAGTTTTTCACGACACTTCCAAGGACGCCATAC
ELI4 ←

AlaCysPheValProGlnEndEnd

183 GCTTGTTTCGTTCCACAATAATAG

CGAACAAAGCAAGGTGTTATTATCCTAG 210

ELI6 ←



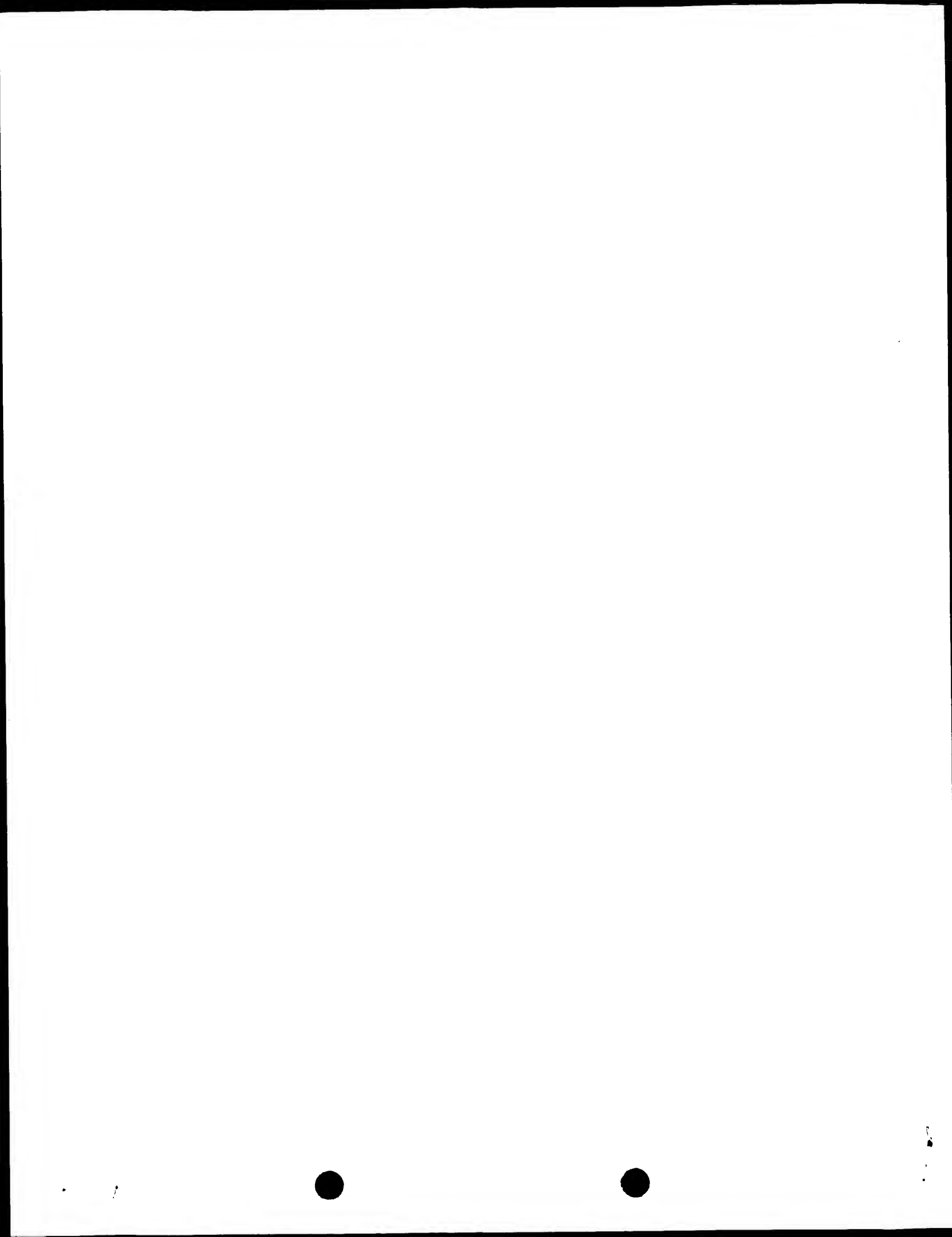
Sun Feb 16 10:28:26 2003

US09833799.pep

Page 1

; Entered [jdelaval 11-Feb-03 8:13]
09-833799-138
agepvkpvstkpqscpiilircamlnppnrcldtdcpgikkceegscgmacfvpq1

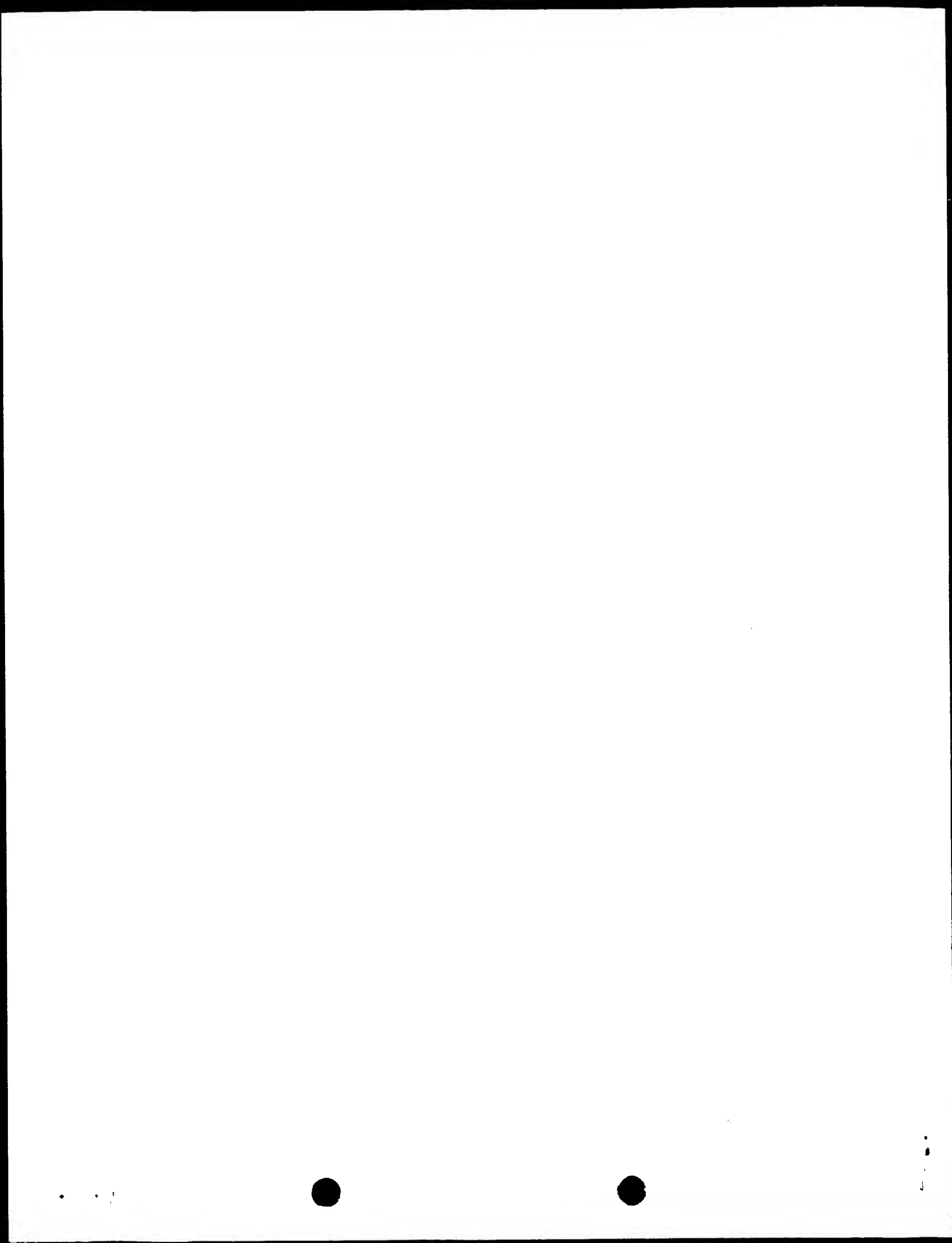
Jan. Delaval
Reference Librarian
Biotechnology & Chemical Library
RM1 1E07 - 703-308-4498
jan.delaval@uspto.gov



; Entered [jdelaval 11-Feb-03 9:38]
09-833799-13A
AAATTCGAGCTCGGTACCATACCTGCATATGCTCAAGAACCAAGTTAAAGGTCCTGTGTCTACTAAGCCAGG
TTCTTGTCTCTATTATCTTGAATTCGTTGCGCTATGTTAAACCCACCTAAACCGTTGTTGAAGGACACTGAT
TGTCAGGTATCAAAAAGTGCTGTGAAGGTTCTCGCGGTATGGCTTGTTTCGTTCCACAATAATAG1

; Entered [jdelaval 11-Feb-03 9:55]
09-833799-13C
gctcgagccatggtatggacgtatacgagttcttgggtcaatttcaggacacacagatgattcgggtccaaga
acaggataatagaactaagcaacgcgatacaatttgggtggattggcaacaaacttcctgtgactaacag
gtccatagtttttcacgacacttccaaggacgccataccgacaagcaagggtgtattatcctag1

Jean Delaval
Science Librarian
Biology & Chemical Library
F07 - 703-308-4498
jdelaval@uspto.gov



GenCore version 5.1.3
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: February 11, 2003, 08:35:56 ; Search time 10 Seconds
(without alignments)
236.415 Million cell updates/sec

Title: 09-833799-13B
Perfect score: 327
Sequence: 1 aqepvkgpvstkgpscpil.....cpgikkcccgsgmacfvpq 57

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 112892 seqs, 41476328 residues

Total number of hits satisfying chosen parameters: 112892

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 100 summaries

Database : SwissProt_40.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	327	100.0	117	1	ELAF_HUMAN
2	221.5	67.7	167	1	ELAF_PIG
3	179.5	54.9	187	1	SPAI_PIG
4	165	50.5	115	1	ALK1_PIG
5	161	49.2	131	1	ALK1_MOUSE
6	160	48.9	144	1	WAP3_PIG
7	158	48.3	55	1	CALU_CAVPO
8	143	43.7	132	1	ALK1_HUMAN
9	101.5	31.0	74	1	WDNM_RAT
10	100.5	30.7	123	1	WFD2_RABIT
11	100.5	30.7	191	1	WAP_MACEU
12	100.5	30.7	676	1	KALM_CHICK
13	98	30.0	137	1	WAP_RAT
14	96.5	29.5	74	1	WDNM_MOUSE
15	95.5	29.2	132	1	WAP_PIG
16	93	28.4	124	1	WFD2_HUMAN
17	92.5	28.3	680	1	KALM_HUMAN
18	90.5	27.7	117	1	WAP_CAMDR
19	88	26.9	124	1	WFD2_CANFA
20	88	26.9	134	1	EPPI_MOUSE
21	85	26.0	133	1	EPPI_HUMAN
22	85	26.0	133	1	EPPI_MACMU
23	81.5	24.9	212	1	WFD1_RAT
24	80.5	24.6	211	1	WFD1_MOUSE
25	79.5	24.3	134	1	WAP_MOUSE
26	79.5	24.3	220	1	WFD1_HUMAN
27	78	23.9	127	1	WAP_RABIT
28	74.5	22.8	110	1	IBP_CARCR
29	67	20.5	34	1	ITR1_MOMCO
30	66	20.2	2911	1	FBN2_HUMAN
31	65	19.9	1700	1	BAR3_CHITE
32	64	19.6	30	1	ITR3_MOMCO
33	64	19.6	34	1	ITR2_MOMCO

34	63	19.3	2907	1	FBN2_MOUSE	Q61555	mus musculu
35	62.5	19.1	202	1	ADEN_ADEB7	P19151	bovine aden
36	62.5	19.1	1964	1	NTC4_MOUSE	P31695	mus musculu
37	62	19.0	71	1	MT2A_MYTED	P80253	mytilus edu
38	62	19.0	71	1	MT2B_MYTED	P80258	mytilus edu
39	61.5	18.8	201	1	ADEN_ADEO7	Q83906	ovine adeno
40	61.5	18.8	1363	1	ILPR_BRALA	O02466	branchiosto
41	61	18.7	129	1	TFF2_HUMAN	Q03403	homo sapien
42	61	18.7	1246	1	YMV2_CAEEL	P34504	caenorhabdi
43	60	18.3	30	1	ITR1_MOMCH	P10294	momordica c
44	60	18.3	71	1	MT21_MYTED	P80251	mytilus edu
45	60	18.3	71	1	MT22_MYTED	P80252	mytilus edu
46	60	18.3	5376	1	ZAN_MOUSE	O88799	mus musculu
47	59.5	18.2	75	1	MT_EISFO	P81695	eisenia foe
48	59.5	18.2	621	1	YC92_CAEEL	P55115	caenorhabdi
49	59.5	18.2	2003	1	NTC4_HUMAN	Q99466	homo sapien
50	59	18.0	2871	1	FBN1_MOUSE	Q61554	mus musculu
51	59	18.0	4289	1	TENX_HUMAN	P22105	homo sapien
52	58	17.7	129	1	TFF2_MOUSE	Q03404	mus musculu
53	58	17.7	694	1	ACSA_CRYPV	Q27549	cryptospori
54	58	17.7	2214	1	SORL_HUMAN	Q92673	h sortilin-
55	58	17.7	2813	1	VWF_CANFA	Q28295	canis famil
56	57.5	17.6	74	1	MT_CRANI	P23038	crassostrea
57	57.5	17.6	204	1	ADEN_ADEF1	Q91ih4	frog adenov
58	57.5	17.6	4660	1	LRP2_RAT	P98158	rattus norv
59	57	17.4	954	1	M3KA_HUMAN	Q02779	homo sapien
60	56.5	17.3	28	1	IEL1_MOMCH	P10296	momordica c
61	56.5	17.3	61	1	MT1B_HUMAN	P07438	homo sapien
62	56.5	17.3	185	1	RL25_CHLTR	O84805	chlamydia t
63	56.5	17.3	326	1	YAW5_SCHPO	Q10180	schizosacch
64	56	17.1	70	1	CX2X_CONBE	Q9u323	conus betul
65	56	17.1	241	1	TR18_HUMAN	Q9y5u5	homo sapien
66	56	17.1	1342	1	ERB3_HUMAN	P21860	homo sapien
67	56	17.1	1680	1	FUR2_DROME	P30432	drosophila
68	56	17.1	2871	1	FBN1_HUMAN	P35555	homo sapien
69	55.5	17.0	78	1	HSTN_VIBCH	P04429	vibrio chol
70	55.5	17.0	78	1	HSTO_VIBCH	Q07425	vibrio chol
71	55.5	17.0	129	1	TFF2_RAT	Q09030	rattus norv
72	55.5	17.0	287	1	TMEP_HUMAN	Q969w9	homo sapien
73	55.5	17.0	500	1	U7I3_HUMAN	Q9bym8	homo sapien
74	55.5	17.0	518	1	VG47_HSVI1	Q00139	ictaluriid h
75	55	16.8	32	1	ITR4_CUCMA	P07853	cucurbita m
76	55	16.8	46	1	EN22_HORSE	P56928	equus cabal
77	55	16.8	855	1	STL4_HUMAN	Q9y5y6	homo sapien
78	55	16.8	1207	1	EGF_HUMAN	P01133	homo sapien
79	54.5	16.7	28	1	ITRA_MOMCH	P30709	momordica c
80	54.5	16.7	329	1	APPF_BACSU	P42065	bacillus su
81	54.5	16.7	462	1	RD21_ARATH	P43297	arabidopsis
82	54.5	16.7	559	1	FZD5_XENLA	P58421	xenopus lae
83	54.5	16.7	684	1	ACS2_KLULA	Q9y7b5	kluveromyc
84	54	16.5	65	1	MTB_STRPU	Q27287	strongyloce
85	54	16.5	127	1	TFF2_PIG	P01359	sus scrofa
86	54	16.5	327	1	Y627_CHLTR	O84632	chlamydia t
87	54	16.5	625	1	XYNA_PIRSP	Q12667	piromyces s
88	54	16.5	769	1	ITB2_BOVIN	P32592	bos taurus
89	54	16.5	1339	1	ERB3_RAT	Q62799	rattus norv
90	54	16.5	2703	1	NOTC_DROME	P07207	drosophila
91	54	16.5	2871	1	FBN1_BOVIN	P98133	bos taurus
92	54	16.5	2871	1	FBN1_PIG	Q9tv36	sus scrofa
93	54	16.5	3767	1	MUA3_CAEEL	P34576	caenorhabdi
94	54	16.5	4753	1	LRP_CAEEL	Q04833	caenorhabdi
95	53.5	16.4	157	1	NEU1_ONCMA	Q07663	oncorhynchu
96	53.5	16.4	328	1	CL70_GIALA	P15799	giardia lam
97	53.5	16.4	346	1	CYSL_LYCES	P20721	lycopersico
98	53.5	16.4	451	1	DADR_XENLA	P42289	xenopus lae
99	53.5	16.4	478	1	DISR_AGRH	P30403	agkistrodon
100	53.5	16.4	604	1	CFAI_RAT	Q9wuw3	rattus norv

ALIGNMENTS

ELAF_HUMAN
ID ELAF_HUMAN STANDARD; PRT; 117 AA.
AC P19957;
DT 01-FEB-1991 (Rel. 17, Created)
DT 01-DEC-1992 (Rel. 24, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Elafin precursor (Elastase-specific inhibitor) (ESI) (Skin-derived
DE antileukoproteinase) (SKALP).
GN PI3.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
[1]
SEQUENCE FROM N.A.
RX MEDLINE=92287100; PubMed=1339270;
RA Saeki T., Ito F., Hagiwara H., Saito Y., Kuroki J., Tachibana S.,
RA Hirose S.;
RT "Primary structure of the human elafin precursor preproelafin deduced
RT from the nucleotide sequence of its gene and the presence of unique
RT repetitive sequences in the prosegment.";
RL Biochem. Biophys. Res. Commun. 185:240-245(1992).
[2]
SEQUENCE FROM N.A., AND SEQUENCE OF 23-48 AND 85-100.
RX TISSUE=Keratinocytes;
RA MEDLINE=93280175; PubMed=7685029;
RA Molhuizen H.O.F., Alkemade H.A.C., Zeeuwen P.L.J.M.,
RA de Jongh G.J., Wieringa B., Schalkwijk J.;
RT "SKALP/elafin: an elastase inhibitor from cultured human
RT keratinocytes. Purification, cDNA sequence, and evidence for
RT transglutaminase cross-linking.";
RL J. Biol. Chem. 268:12028-12032(1993).
[3]
SEQUENCE FROM N.A.
RX TISSUE=Placenta;
RA MEDLINE=93236929; PubMed=8476637;
RA Sallenave J.M., Silva A.;
RT "Characterization and gene sequence of the precursor of elafin, an
RT elastase-specific inhibitor in bronchial secretions.";
RL Am. J. Respir. Cell Mol. Biol. 8:439-445(1993).
[4]
SEQUENCE FROM N.A.
RX MEDLINE=21638749; PubMed=11780052;
RA Deloukas P., Matthews L.H., Ashurst J., Burton J., Gilbert J.G.R.,
RA Jones M., Stavrides G., Almeida J.P., Babbage A.K., Bagguely C.L.,
RA Bailey J., Barlow K.F., Bates K.N., Beard L.M., Beare D.M.,
RA Beasley O.P., Bird C.P., Blakey S.E., Bridgeman A.M., Brown A.J.,
RA Buck D., Burrill W.D., Butler A.P., Carder C., Carter N.P.,
RA Chapman J.C., Clamp G., Clark G., Clark L.N., Clark S.Y., Clee C.M.,
RA Clegg S., Cobley V.E., Collier R.E., Connor R.E., Corby N.R.,
RA Coulson A., Coville G.J., Deadman R., Dhami P.D., Dunn M.,
RA Ellington A.G., Frankland J.A., Fraser A., French L., Garner P.,
RA Grafham D.V., Griffiths C., Griffiths M.N.D., Gwilliam R., Hall R.E.,
RA Hammond S., Harley J.L., Heath P.D., Ho S., Holden J.L., Howden P.J.,
RA Huckle E., Hunt A.R., Hunt S.E., Jekosch K., Johnson C.M., Johnson D.,
RA Kay M.P., Kimberley A.M., King A., Knights A., Laird G.K., Lawlor S.,
RA Lehtvaslao M.H., Leversha M.A., Lloyd C., Lloyd D.M., Lovell J.D.,
RA Marsh V.L., Martin S.L., McConnachie L.J., McMay K., McMurray A.A.,
RA Milne S.A., Mistry D., Moore M.J.F., Mullikin J.C., Nickerson T.,
RA Oliver K., Parker A., Patel R., Pearce T.A.V., Peck A.I.,
RA Phillimore B.J.C.T., Prathalingam S.R., Plumb R.W., Ramsay H.,
RA Rice C.M., Ross M.T., Scott C.E., Sehra H.K., Showkneen R., Sims S.,
RA Skuce C.D., Smith M.L., Soderlund C., Steward C.A., Sulston J.E.,
RA Swann R.M., Sycamore N., Taylor R., Tee L., Thomas D.W., Thorpe A.,
RA Tracey A., Tromans A.C., Vaudin M., Wall M., Wallis J.M.,
RA Whitehead S.L., Whittaker P., Willey D.L., Williams L., Williams S.A.,
RA Wilming L., Wray P.W., Hubbard T., Durbin R.M., Bentley D.R., Beck S.,
RA Rogers J.;
RT "The DNA sequence and comparative analysis of human chromosome 20.";
RL Nature 414:865-871(2001).
[5]
SEQUENCE OF 61-117.
RX TISSUE=Horny layers;

RX MEDLINE=90368643; PubMed=2394696;
RA Wiedow O., Schroeder J.-M., Gregory H., Young J.A., Christophers E.;
RT "Elafin: an elastase-specific inhibitor of human skin. Purification,
RT characterization, and complete amino acid sequence.";
RL J. Biol. Chem. 265:14791-14795(1990).
[6]
ERRATUM.
RA Wiedow O., Schroeder J.-M., Gregory H., Young J.A., Christophers E.;
RA J. Biol. Chem. 266:3356-3356(1991).
[7]
SEQUENCE OF 70-94.
RX TISSUE=Sputum;
RA MEDLINE=91248412; PubMed=2039600;
RA Sallenave J.-M., Ryle A.P.;
RT "Purification and characterization of elastase-specific inhibitor.
RT Sequence homology with mucus proteinase inhibitor.";
RL Biol. Chem. Hoppe-Seyler 372:13-21(1991).
[8]
SEQUENCE OF 61-78.
RX TISSUE=Sputum;
RA MEDLINE=92162196; PubMed=1536690;
RA Sallenave J.-M., Marsden M.D., Ryle A.P.;
RT "Isolation of elafin and elastase-specific inhibitor (ESI) from
RT bronchial secretions. Evidence of sequence homology and immunological
RT cross-reactivity.";
RL Biol. Chem. Hoppe-Seyler 373:27-33(1992).
[9]
SEQUENCE OF 85-100.
RX TISSUE=Keratinocytes;
RA MEDLINE=91159523; PubMed=2001428;
RA Schalkwijk J., de Roo C., de Jongh G.J.;
RT "Skin-derived antileukoproteinase (SKALP), an elastase inhibitor from
RT human keratinocytes. Purification and biochemical properties.";
RL Biochim. Biophys. Acta 1096:148-154(1991).
[10]
X-RAY CRYSTALLOGRAPHY (1.9 ANGSTROMS).
RX MEDLINE=96387196; PubMed=8794736;
RA Tsunemi M., Matsuura Y., Sakakibara S., Katsube Y.;
RT "Crystal structure of an elastase-specific inhibitor elafin complexed
RT with porcine pancreatic elastase determined at 1.9-A resolution.";
RL Biochemistry 35:11570-11576(1996).
[11]
STRUCTURE BY NMR.
RX MEDLINE=97315116; PubMed=9171290;
RA Francart C., Dauchez M., Alix A.J.P., Lippens G.;
RT "Solution structure of R-elafin, a specific inhibitor of elastase.";
RL J. Mol. Biol. 268:666-677(1997).
CC -!- FUNCTION: NEUTROPHIL AND PANCREATIC ELASTASE-SPECIFIC INHIBITOR OF
CC SKIN. IT MAY PREVENT ELASTASE-MEDIATED TISSUE PROTEOLYSIS.
CC -!- SUBCELLULAR LOCATION: Secreted.
CC -!- DOMAIN: CONSISTS OF TWO DOMAINS: THE TRANSGLUTAMINASE SUBSTRATE
CC DOMAIN (CEMENTOIN MOIETY) AND THE ELASTASE INHIBITOR DOMAIN. THE
CC TRANSGLUTAMINASE SUBSTRATE DOMAIN SERVES AS AN ANCHOR TO LOCALIZE
CC ELAFIN COVALENTLY TO SPECIFIC SITES ON EXTRACELLULAR MATRIX
CC PROTEINS.
CC -!- SIMILARITY: CONTAINS 1 WAP-TYPE DOMAIN.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).

CC EMBL; Z18538; CAA79223.1; -.
CC EMBL; D13156; BAA02441.1; -.
CC EMBL; S58717; AAB26371.1; -.
CC EMBL; L10343; AAA36483.1; -.
CC EMBL; AL049767; CAB53524.1; -.
CC PIR; JH0614; JH0614.
CC PIR; S13607; S13607.
CC PIR; A46749; A46749.

```

DR InterPro; IPR002221; WAP.
DR Pfam; PF00095; wap; 1.
DR PRINTS; PR00003; 4DISULPHCORE.
DR SMART; SM00217; WAP; 1.
DR PROSITE; PS00317; 4 DISULFIDE CORE; 1.
KW Serine protease inhibitor; Signal; Repeat.
FT SIGNAL 1 21 POTENTIAL.
FT PROPEP 22 70 POTENTIAL.
FT CHAIN 71 167 ELAFIN.
FT DOMAIN 44 115 12 X 6 AA APPROXIMATE TANDEM REPEATS.
FT REPEAT 44 49 1.
FT REPEAT 50 55 2.
FT REPEAT 56 61 3.
FT REPEAT 62 67 4.
FT REPEAT 68 73 5.
FT REPEAT 74 79 6.
FT REPEAT 80 85 7.
FT REPEAT 86 91 8.
FT REPEAT 92 97 9.
FT REPEAT 98 103 10.
FT REPEAT 104 109 11.
FT REPEAT 110 115 12.
FT DOMAIN 78 126 2 X TANDEM REPEATS OF SVP-1 LIKE MOTI
FT REPEAT 78 101 SVP-1 CLOTTING 1.
FT REPEAT 102 126 SVP-1 CLOTTING 2.
FT DOMAIN 122 167 WAP.
FT DISULFID 126 155 BY SIMILARITY.
FT DISULFID 133 159 BY SIMILARITY.
FT DISULFID 142 154 BY SIMILARITY.
FT DISULFID 148 163 BY SIMILARITY.
SQ SEQUENCE 167 AA; 17923 MW; 25932EALA459D9CA CRC64;

Query Match 67.7%; Score 221.5; DB 1; Length 167;
Best Local Similarity 61.9%; Pred. No. 1.7e-17;
Matches 39; Conservative 5; Mismatches 12; Indels 7; Gaps 1

QY 2 QEPVKGP-----VSTKPGSCPILIRCAMLNPPNRLKDTDCPGIKKCCGSCGMACF
Db 105 QDPVKAQPAIKRLILLTKPGSCPRIILRLCMVNPNNRCLSDAQCPGLKKCCGFCGKACM
QY 55 VPQ 57
Db 165 DPK 167

RESULT 3
SPAI_PIG
ID SPAI_PIG STANDARD; PRT; 187 AA.
AC P16225;
DT 01-APR-1990 (Rel. 14, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Sodium/potassium ATPase inhibitor SPAI-2 precursor (WAP-2 protein).
OS Sus scrofa (pig).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
OX NCBI_Taxid=9823;
RN [1]
RP SEQUENCE FROM N.A., AND SEQUENCE OF 22-58.
RC TISSUE=Duodenum;
RX MEDLINE=95403443; PubMed=7673229;
RA Kuroki J., Hosoya T., Itakura M., Hirose S., Tamechika I.,
RA Yoshimoto T., Ghoneim M.A., Nara K., Kato A., Suzuki Y., Furukawa M.
RA Tachibana S.;
RT "Cloning, characterization, and tissue distribution of porcine SPAI
RT a protein with a transglutaminase substrate domain and the WAP
RT motif.";
RL J. Biol. Chem. 270:22428-22433(1995).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=96215132; PubMed=8636131;
RA Tamechika I., Itakura M., Saruta Y., Furukawa M., Kato A.,
RA Tachibana S., Hirose S.;

```



```
KW Serine protease inhibitor; Repeat.
FT DOMAIN 8 62 TRYPSIN INHIBITORY DOMAIN.
FT DOMAIN 63 114 ELASTASE INHIBITORY DOMAIN.
FT DOMAIN 14 59 WAP 1.
FT DOMAIN 68 113 WAP 2.
FT ACT_SITE 28 28 TRYPSIN INHIBITORY SITE (PROBABLE).
FT ACT_SITE 81 82 ELASTASE OR CHYMOTRYPSIN INHIBITORY SITE
      (PROBABLE).
FT DISULFID 18 47 BY SIMILARITY.
FT DISULFID 26 51 BY SIMILARITY.
FT DISULFID 34 46 BY SIMILARITY.
FT DISULFID 40 55 BY SIMILARITY.
FT DISULFID 72 101 BY SIMILARITY.
FT DISULFID 79 105 BY SIMILARITY.
FT DISULFID 88 100 BY SIMILARITY.
FT DISULFID 94 109 BY SIMILARITY.
SQ SEQUENCE 115 AA; C8C937E997815BCF CRC64;

Query Match 50.5%; Score 165; DB 1; Length 115;
Best Local Similarity 48.1%; Pred. No. 1.7e-11;
Matches 25; Conservative 6; Mismatches 21; Indels 0; Gaps 0;

QY 5 VKGPVSTKPGSCPILIRCAMLNPPNRCCLKDTCPGIKKCEGSCGMACFVP 56
Db 61 ITNPVKVKGKCPVVGQCMWMLNPPNHCKTDSQCLGDLKCKSMCGKVCLTP 112

RESULT 5
ALK1_MOUSE
ID ALK1_MOUSE STANDARD; PRT; 131 AA.
AC P97430; O09081; O09082;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Antileukoproteinase 1 precursor (ALP) (Secretory leukocyte protease
  inhibitor).
GN SLPI.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
[1] SEQUENCE FROM N.A.
RX MEDLINE=97191310; PubMed=9039268;
RA Jin F.-Y., Nathan C.F., Radzioch D., Ding A.;
RT "Secretory leukocyte protease inhibitor: a macrophage product induced
  by and antagonistic to bacterial lipopolysaccharide.";
RL Cell 88:417-426(1997).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=97271386; PubMed=9126337;
RA Zitnik R.J., Zhang J., Kashem M.A., Kohno T., Lyons D.E., Wright C.D.,
RA Rosen E., Goldberg I., Hayday A.C.;
RT "The cloning and characterization of a murine secretory leukocyte
  protease inhibitor cDNA.";
RL Biochem. Biophys. Res. Commun. 232:687-697(1997).
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6 X CBA; TISSUE=Lung;
RX MEDLINE=98011992; PubMed=9351627;
RA Abe T., Tominaga Y., Kikuchi T., Watanabe A., Satoh K., Watanabe Y.,
RA Nukiwa T.;
RT "Bacterial pneumonia causes augmented expression of the secretory
  leukoprotease inhibitor gene in the murine lung.";
RL Am. J. Respir. Crit. Care Med. 156:1235-1240(1997).
CC -!- FUNCTION: ACID-STABLE PROTEINASE INHIBITOR WITH STRONG AFFINITIES
  FOR TRYPSIN, CHYMOTRYPSIN, ELASTASE, AND CATHEPSIN G. MAY PREVENT
  ELASTASE-MEDIATED DAMAGE TO ORAL AND POSSIBLY OTHER MUCOSAL
  TISSUES (BY SIMILARITY).
CC -!- TISSUE SPECIFICITY: HIGHEST EXPRESSION IN LUNG, SPLEEN,
  INTESTINE AND EPIDIDYMIS WITH LOWER LEVELS IN LIVER AND SEMINAL
  VESICLE. NO EXPRESSION IN BRAIN, HEART, KIDNEY AND MUSCLE.
CC -!- SIMILARITY: CONTAINS 2 WAP-TYPE DOMAINS.
```

```
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; U73004; AAC53047.1; -.
DR EMBL; U88093; AAC53140.1; -.
DR EMBL; U94341; AAC53394.1; -.
DR HSSP; P19957; 2REL.
DR MGD; MGI:109297; Slpi.
DR InterPro; IPR002221; WAP.
DR Pfam; PF00095; wap; 2.
DR ProDom; PD001224; WAP; 1.
DR SMART; SM00217; WAP; 2.
DR PROSITE; PS00317; 4_DISULFIDE_CORE; 2.
KW Serine protease inhibitor; Repeat; Signal.
FT SIGNAL 1 25 BY SIMILARITY.
FT CHAIN 26 131 ANTILEUKOPROTEINASE 1.
FT DOMAIN 26 84 TRYPSIN INHIBITORY DOMAIN.
FT DOMAIN 85 131 ELASTASE INHIBITORY DOMAIN.
FT DOMAIN 32 77 WAP 1.
FT DOMAIN 86 131 WAP 2.
FT ACT_SITE 46 46 TRYPSIN INHIBITORY SITE (PROBABLE).
FT ACT_SITE 98 99 ELASTASE OR CHYMOTRYPSIN INHIBITORY SITE
      (PROBABLE).
FT DISULFID 36 65 BY SIMILARITY.
FT DISULFID 44 69 BY SIMILARITY.
FT DISULFID 52 64 BY SIMILARITY.
FT DISULFID 58 73 BY SIMILARITY.
FT DISULFID 90 119 BY SIMILARITY.
FT DISULFID 97 123 BY SIMILARITY.
FT DISULFID 106 118 BY SIMILARITY.
FT DISULFID 112 127 BY SIMILARITY.
SQ SEQUENCE 131 AA; 14308 MW; A57C9E30FE711B8F CRC64;

Query Match 49.2%; Score 161; DB 1; Length 131;
Best Local Similarity 52.8%; Pred. No. 5.1e-11;
Matches 28; Conservative 3; Mismatches 22; Indels 0; Gaps 0;

QY 4 PVKGPVSTKPGSCPILIRCAMLNPPNRCCLKDTCPGIKKCEGSCGMACFVP 56
Db 78 PIRKPVWRKPGRCVKVTQARCMMLNPPNVQORDGQCDGKYKCEGICGKVCLPP 130

RESULT 6
WAP3_PIG
ID WAP3_PIG STANDARD; PRT; 144 AA.
AC Q29126;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE WAP-3 protein precursor.
OS Sus scrofa (Pig).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
OX NCBI_TaxID=9823;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=96215132; PubMed=8636131;
RA Tamechika I., Itakura M., Saruta Y., Furukawa M., Kato A.,
RA Tachibana S., Hirose S.;
RT "Accelerated evolution in inhibitor domains of porcine elafin family
  members.";
RL J. Biol. Chem. 271:7012-7018(1996).
CC -!- TISSUE SPECIFICITY: LARGE INTESTINE (RELATIVELY LOW LEVELS).
CC -!- SIMILARITY: CONTAINS 1 WAP-TYPE DOMAIN.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
```


RA Dear T.N., Kefford R.F.;

RT "The WDNM1 gene product is a novel member of the 'four-disulphide

RL core' family of proteins.";

RL Biochem. Biophys. Res. Commun. 176:247-254(1991).

CC -!- FUNCTION: INVOLVED IN THE METASTATIC POTENTIAL OF ADENOCARCINOMAS

CC -!- IN RAT: COULD HAVE PROTEINASE INHIBITING CAPACITY.

CC -!- TISSUE SPECIFICITY: SPLEEN, KIDNEY, BRAIN, LIVER, LUNG, AND CELL

CC LINE RAT-2.

CC -!- SIMILARITY: CONTAINS 1 WAP-TYPE DOMAIN.

CC

CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

CC the European Bioinformatics Institute. There are no restrictions on its

CC use by non-profit institutions as long as its content is in no way

CC modified and this statement is not removed. Usage by and for commercial

CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>

CC or send an email to license@isb-sib.ch).

CC

DR EMBL; X13309; CAA31688.1; ALT_INIT.

DR PIR; JH0390; JH0390.

DR HSSP; P19957; 2REL.

DR InterPro; IPR002221; WAP.

DR Pfam; PF00095; wap; 1.

DR SMART; SM00217; WAP; 1.

DR PROSITE; PS00317; 4_DISULFIDE_CORE; 1.

KW Protease inhibitor; Signal.

FT NON_TER 1

FT SIGNAL <1 23 POTENTIAL.

FT CHAIN 24 74 WDNM1 PROTEIN.

FT DOMAIN 29 73 WAP.

SQ SEQUENCE 74 AA; 7740 MW; 51802C70CDAF0521 CRC64;

Query Match 31.0%; Score 101.5; DB 1; Length 74;

Best Local Similarity 38.5%; Pred. No. 9.3e-05;

Matches 20; Conservative 1; Mismatches 16; Indels 15; Gaps 2;

QY 12 KPGSCPIILIRCAMLNPPNR-----CLKDTCPGIKKCCGSCGMACFVP 56

DB 29 KPGKCP-----KNPPRSIGTCVELCGDQSCPNIQKCCNSGCHVCKSP 72

RESULT 10

WFD2_RABIT

ID WFD2_RABIT STANDARD; PRT; 123 AA.

AC Q28631;

DT 15-JUL-1998 (Rel. 36, Created)

DT 16-OCT-2001 (Rel. 40, Last sequence update)

DT 15-JUN-2002 (Rel. 41, Last annotation update)

DE Major epididymis-specific protein E4 precursor (Epididymal protein BE-

DE 20) (WAP four-disulphide core domain protein 2).

GN WFD2.

OS Oryctolagus cuniculus (Rabbit).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.

OX NCHI_TaxID=9986;

RN [1]

RP SEQUENCE FROM N.A., AND REVISION TO 29.

RC TISSUE=Epididymis;

RX MEDLINE=99201788; PubMed=10101572;

RA Fan H.-Y., Miao S.-Y., Wang L.-F., Koide S.S.;

RT "Expression and characterization of an epididymis-specific gene.";

RL Arch. Androl. 42:63-69(1999).

RN [2]

RP SEQUENCE OF 29-123 FROM N.A., AND SEQUENCE OF 29-58.

RC TISSUE=Epididymis;

RX MEDLINE=97040973; PubMed=8886263;

RA Xu W.D., Wang L.-F., Miao S.-Y., Zhao M., Fan H.-Y., Zong S.D.,

RA Wu Y.W., Shi X.Q., Koide S.S.;

RT "Identification of a rabbit epididymal protein gene.";

RL Arch. Androl. 37:135-141(1996).

CC -!- FUNCTION: POSSIBLE FUNCTION IN SPERM MATURATION.

CC -!- SUBCELLULAR LOCATION: Secreted.

CC -!- TISSUE SPECIFICITY: EPIDIDYMIS.

CC -!- SIMILARITY: CONTAINS 2 WAP-TYPE DOMAINS.

CC

CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

CC the European Bioinformatics Institute. There are no restrictions on its

CC use by non-profit institutions as long as its content is in no way

CC modified and this statement is not removed. Usage by and for commercial

CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>

CC or send an email to license@isb-sib.ch).

CC

DR EMBL; U26725; AAA66525.2; -.

DR HSSP; Q9N0L8; 1TWP.

DR InterPro; IPR002221; WAP.

DR Pfam; PF00095; wap; 2.

DR PRINTS; PR00003; 4DISULPHCORE.

DR ProDom; PD001224; WAP; 1.

DR SMART; SM00217; WAP; 2.

DR PROSITE; PS00317; 4_DISULFIDE_CORE; 2.

KW Repeat; Signal; Glycoprotein.

FT SIGNAL 1 28 MAJOR EPIDIDYMIS-SPECIFIC PROTEIN E4.

FT CHAIN 29 123 WAP 1.

FT DOMAIN 33 75 WAP 2.

FT DOMAIN 77 123

FT DISULFID 37 63 BY SIMILARITY.

FT DISULFID 46 67 BY SIMILARITY.

FT DISULFID 50 62 BY SIMILARITY.

FT DISULFID 56 71 BY SIMILARITY.

FT DISULFID 81 109 BY SIMILARITY.

FT DISULFID 92 113 BY SIMILARITY.

FT DISULFID 96 108 BY SIMILARITY.

FT DISULFID 102 118 BY SIMILARITY.

FT CARBOHYD 45 45 N-LINKED (GLCNAC...) (POTENTIAL).

SQ SEQUENCE 123 AA; 12803 MW; 99F9A649CCBB5B32 CRC64;

Query Match 30.7%; Score 100.5; DB 1; Length 123;

Best Local Similarity 39.6%; Pred. No. 0.00018;

Matches 21; Conservative 3; Mismatches 26; Indels 3; Gaps 1;

QY 4 PVKGPVSTKPGSCPIILIRCAMLNPPNRCLKDTDCPGIKKCCGSCGMACFVP 56

DB 25 PVTGTGADKPGVCPQL---SADLNCTQDCRADQDCAENLKCCRCAGCSAICSIP 74

RESULT 11

WAP_MACEU

ID WAP_MACEU STANDARD; PRT; 191 AA.

AC Q9N0L8;

DT 16-OCT-2001 (Rel. 40, Created)

DT 16-OCT-2001 (Rel. 40, Last sequence update)

DT 15-JUN-2002 (Rel. 41, Last annotation update)

DE Whey acidic protein precursor (tWAP).

GN WAP.

OS Macropus eugenii (Tamar wallaby).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Metatheria; Diprotodontia; Macropodidae; Macropus.

OX NCHI_TaxID=9315;

RN [1]

RP SEQUENCE FROM N.A., AND 3D-STRUCTURE MODELING.

RC TISSUE=Lactating mammary gland;

RX MEDLINE=20390063; PubMed=10801834;

RA Simpson K.J., Ranganathan S., Fisher J.A., Janssens P.A., Shaw D.C.,

RA Nicholas K.R.;

RT "The gene for a novel member of the whey acidic protein family encodes

RT three four-disulphide core domains and is asynchronously expressed

RT during lactation.";

RL J. Biol. Chem. 275:23074-23081(2000).

CC -!- FUNCTION: COULD BE A PROTEASE INHIBITOR.

CC -!- SUBCELLULAR LOCATION: Secreted.

CC -!- TISSUE SPECIFICITY: MILK-SPECIFIC; MAJOR PROTEIN COMPONENT OF MILK

CC WHEY.

CC -!- SIMILARITY: CONTAINS 3 WAP-TYPE DOMAINS.

CC

CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----

DR EMBL; AJ005356; CAB90357.1; -.
DR PDB; 1TWP; 13-JAN-00.
DR InterPro; IPR002221; WAP.
DR Pfam; PF00095; wap; 3.
DR PRINTS; PR00003; 4DISULPHCORE.
DR ProDom; PD001224; WAP; 1.
DR SMART; SM00217; WAP; 3.
DR PROSITE; PS00317; 4 DISULFIDE CORE; 3.
KW Milk; Whey; Protease inhibitor; Repeat; Signal; 3D-structure.
FT SIGNAL 1 18 POTENTIAL.
FT CHAIN 19 191 WHEY ACIDIC PROTEIN.
FT DOMAIN 25 71 WAP 1.
FT DOMAIN 74 126 WAP 2.
FT DOMAIN 129 175 WAP 3.
FT DISULFID 29 59
FT DISULFID 42 63
FT DISULFID 46 58
FT DISULFID 52 67
FT DISULFID 78 114
FT DISULFID 97 118
FT DISULFID 101 113
FT DISULFID 107 122
FT DISULFID 133 163
FT DISULFID 140 167
FT DISULFID 150 162
FT DISULFID 156 171
FT CARBOHYD 69
SQ SEQUENCE 191 AA; 21133 MW; BF90B81DADBBE50D CRC64; N-LINKED (GLCNAC. . .) (POTENTIAL).
Query Match 30.7%; Score 100.5; DB 1; Length 191;
Best Local Similarity 37.0%; Pred. No. 0.00026;
Matches 20; Conservative 4; Mismatches 21; Indels 9; Gaps 2;

QY 8 PVSTKPGSCPILIRCAMLNPPNR-----CLKDTCPCGIKKCCGSCGMACFVP 56
Db 125 PVKAKPGRCPAVTGIC----PEKKSWFHTCQRDDQCKENKCCSSACGRRCTNP 174

RESULT 12
KALM_CHICK STANDARD; PRT; 676 AA.
ID KALM_CHICK
AC P33005;
DT 01-OCT-1993 (Rel. 27, Created)
DT 01-OCT-1993 (Rel. 27, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Anosmin 1 precursor (Kallmann syndrome protein homolog).
GN KAL.
OS Gallus gallus (Chicken).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
OC Gallus.
OX NCBI_TaxID=9031;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Brain;
RX MEDLINE=94010957; PubMed=8406507;
RA Legouis R., Cohen-Salmon M., del Castillo I., Levilliers J.,
RA Capy L., Mornon J.-P., Petit C.;
RT "Characterization of the chicken and quail homologues of the human
RT gene responsible for the X-linked Kallmann syndrome.";
RL Genomics 17:516-518(1993).
CC -!- FUNCTION: May be an adhesion-like molecule with anti-protease
CC activity.
CC -!- SUBCELLULAR LOCATION: Secreted. Localized at cell surface (By
CC similarity).
CC -!- SIMILARITY: CONTAINS 4 FIBRONECTIN TYPE III-LIKE DOMAINS.

CC -!- SIMILARITY: CONTAINS 1 WAP-TYPE DOMAIN.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----

DR EMBL; L12144; AAA51435.1; -.
DR PIR; B47222; B47222.
DR HSSP; P19957; 2REL.
DR InterPro; IPR003961; FN_III.
DR InterPro; IPR002221; WAP.
DR Pfam; PF00041; fn3; 3.
DR Pfam; PF00095; wap; 1.
DR PRINTS; PR00003; 4DISULPHCORE.
DR SMART; SM00060; FN3; 3.
DR SMART; SM00217; WAP; 1.
DR PROSITE; PS00317; 4 DISULFIDE CORE; 1.
KW Cell adhesion; Glycoprotein; Serine protease inhibitor; Repeat;
KW Signal.
FT SIGNAL 1 21 POTENTIAL.
FT CHAIN 22 676 ANOSMIN 1.
FT DOMAIN 22 115 "CYSTEINE BOX".
FT DOMAIN 125 171 WAP.
FT DOMAIN 176 280 FIBRONECTIN TYPE-III 1.
FT DOMAIN 281 397 FIBRONECTIN TYPE-III 2.
FT DOMAIN 398 536 FIBRONECTIN TYPE-III 3.
FT DOMAIN 537 657 FIBRONECTIN TYPE-III 4.
FT CARBOHYD 66 66 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 204 204 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 295 295 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 466 466 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 549 549 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 560 560 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 676 AA; 76375 MW; 3FAC7ED82EA7E352 CRC64;

Query Match 30.7%; Score 100.5; DB 1; Length 676;
Best Local Similarity 34.4%; Pred. No. 0.00074;
Matches 21; Conservative 6; Mismatches 23; Indels 11; Gaps 2;

QY 3 EPVKGPVSTKPGSCPI-----ILIRCAMLNPPNRCLKDTCPCGIKKCCGSCGMACFVP 56
Db 116 EFLKYILSVKQGDPCPAPEKASGFAAACV-----ESCEADSECSGVKKCCSNGCGHTCQVP 170

QY 57 Q 57
Db 171 K 171
RESULT 13
WAP_RAT STANDARD; PRT; 137 AA.
ID WAP_RAT
AC P01174;
DT 21-JUL-1986 (Rel. 01, Created)
DT 01-OCT-1989 (Rel. 12, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Whey acidic protein precursor (Whey phosphoprotein) (WAP).
GN WAP.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=85062841; PubMed=6095207;
RA Campbell S.M., Rosen J.M., Hennighausen L.G., Strech-Jurk U.,
RA Sippel A.E.;
RT "Comparison of the whey acidic protein genes of the rat and mouse.";
RL Nucleic Acids Res. 12:8685-8697(1984).
RN [2]


```
CC -----
DR EMBL; X63187; CAA44869.1; -.
DR EMBL; A18924; CAA01433.1; -.
DR EMBL; AL031663; CAB37641.1; -.
DR HSSP; Q9N0L8; 1TWP.
DR Genew; HGNC:15939; WFDC2.
DR InterPro; IPR002221; WAP.
DR Pfam; PF00095; wap; 3.
DR SMART; SM00217; WAP; 1.
DR PROSITE; PS00317; 4_DISULFIDE_CORE; 2.
KW Repeat; Signal; Glycoprotein.
FT SIGNAL 1 21 POTENTIAL.
FT CHAIN 22 124 MAJOR EPIDIDYMIS-SPECIFIC PROTEIN E4.
FT DOMAIN 32 73 WAP 1.
FT DOMAIN 76 124 WAP 2.
FT DISULFID 36 62 BY SIMILARITY.
FT DISULFID 45 66 BY SIMILARITY.
FT DISULFID 49 61 BY SIMILARITY.
FT DISULFID 55 70 BY SIMILARITY.
FT DISULFID 80 110 BY SIMILARITY.
FT DISULFID 93 114 BY SIMILARITY.
FT DISULFID 97 109 BY SIMILARITY.
FT DISULFID 103 119 BY SIMILARITY.
FT CARBOHYD 44 44 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CONFLICT 71 72 SL -> LLC (IN REF. 1).
FT CONFLICT 101 101 S -> T (IN REF. 1).
FT SEQUENCE 124 AA; 12993 MW; 9536B00B385259AD CRC64;
SQ
Query Match 28.4%; Score 93; DB 1; Length 124;
Best Local Similarity 40.4%; Pred. No. 0.0012;
Matches 19; Conservative 6; Mismatches 20; Indels 2; Gaps 2;
OY 12 KPGSCPILLRCAMLN-PPNRCCLKDTPGKIKKCEGSCG-MACFVP 56
Db 76 KEGSCPQVNIFFPQLGLCRDQCQVDSQCPGQMKCCRNCGKVCVTP 122
RESULT 17
KALM HUMAN
ID KALM HUMAN STANDARD; PRT; 680 AA.
AC P23352;
DT 01-NOV-1991 (Rel. 20, Created)
DT 01-DEC-1992 (Rel. 24, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Anosmin 1 precursor (Kallmann syndrome protein) (Adhesion molecule-
DE like X-linked).
GN KAL1 OR KAL OR ADMLX OR KALIG1.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=92005720; PubMed=1913827;
RA Legouis R., Hardelin J.-P., Levilliers J., Claverie J.-M., Compain S.,
RA Wunderle V., Millaudeau P., le Paslier D., Cohen D., Caterina D.,
RA Bougueleret L., Delemarre-Van de Waal H., Lutfalla G., Weissenbach J.,
RA Petit C.;
RT "The candidate gene for the X-linked Kallmann syndrome encodes a
RT protein related to adhesion molecules.";
RL Cell 67:423-435(1991).
RN [2]
RP REVISIONS.
RX MEDLINE=93265164; PubMed=1303284;
RA del Castillo I., Cohen-Salmon M., Blanchard S., Lutfalla G.,
RA Petit C.;
RT "Structure of the X-linked Kallmann syndrome gene and its homologous
RT pseudogene on the Y chromosome.";
RL Nat. Genet. 2:305-310(1992).
RN [3]
RP SEQUENCE FROM N.A.
RX MEDLINE=92018217; PubMed=1922361;
RA Franco B., Guioli S., Pragliola A., Inceri B., Bardoni B.,
```

```
RA Tonlorenzi R., Carrozo R., Maestrini E., Pieretti M.,
RA Taillon-Miller P., Brown C.J., Willard H.F., Lawrence C.,
RA Persico N.G., Camerino G., Ballabio A.;
RT "A gene deleted in Kallmann's syndrome shares homology with neural
RL cell adhesion and axonal path-finding molecules.";
RN Nature 353:529-536(1991).
RN [4]
RP SEQUENCE OF 1-71 FROM N.A.
RX MEDLINE=96069588; PubMed=7590336;
RA Cohen-Salmon M., Tronche F., del Castillo I., Petit C.;
RT "Characterization of the promoter of the human KAL gene, responsible
RT for the X-chromosome-linked Kallmann syndrome.";
RL Gene 164:235-242(1995).
RN [5]
RP CHARACTERIZATION.
RX PubMed=8832397;
RA Soussi-Yanicostas N., Hardelin J.-P., del Mar Arroyo-Jimenez M.,
RA Ardouin O., Legouis R., Levilliers J., Traincard F., Betton J.-M.,
RA Cabanie L., Petit C.;
RT "Initial characterization of anosmin-1, a putative extracellular
RT matrix protein synthesized by definite neuronal cell populations in
RT the central nervous system.";
RL J. Cell Sci. 109:1749-1757(1996).
RN [6]
RP VARIANT KAL1 LYS-267, AND VARIANT VAL-534.
RX MEDLINE=93278384; PubMed=8504298;
RA Hardelin J.-P., Levilliers J., Blanchard S., Carel J.-C.,
RA Leutenegger M., Pinard-Bertelletto J.-P., Bouloux P., Petit C.;
RT "Heterogeneity in the mutations responsible for X chromosome-linked
RT Kallmann syndrome.";
RL Hum. Mol. Genet. 2:373-377(1993).
RN [7]
RP VARIANT KAL1 LYS-514, AND VARIANT VAL-534.
RX MEDLINE=98251583; PubMed=9589672;
RA Maya-Nunez G., Zenteno J.C., Ulloa-Aguirre A., Kofman-Alfaro S.,
RA Mendez J.P.;
RT "A recurrent missense mutation in the KAL gene in patients with
RT X-linked Kallmann's syndrome.";
RL J. Clin. Endocrinol. Metab. 83:1650-1653(1998).
CC -!- FUNCTION: May be an adhesion-like molecule with anti-protease
CC activity.
CC -!- SUBCELLULAR LOCATION: Secreted. Localized at cell surface.
CC -!- PTM: N-glycosylated.
CC -!- DISEASE: Defects in KAL1 are a cause of Kallmann syndrome (KAL1 OR
CC KS), a genetic disorder that associates hypogonadotropic.
CC hypogonadism and anosmia. In this disease, the normal embryonic
CC migration of GNRH-synthesizing neurons from the olfactory placodes
CC to the hypothalamic region as well as the axonal extension of
CC olfactory neurons towards the forebrain are impaired.
CC -!- SIMILARITY: CONTAINS 4 FIBRONECTIN TYPE III-LIKE DOMAINS.
CC -!- SIMILARITY: CONTAINS 1 WAP-TYPE DOMAIN.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; M97252; AAA59202.1; -.
DR EMBL; S60085; AAB20108.1; ALT_SEQ.
DR EMBL; X60299; CAA42841.1; -.
DR EMBL; X82034; CAA57554.1; -.
DR PIR; A40351; A40351.
DR PIR; S17982; S17982.
DR Genew; HGNC:6211; KAL1.
DR MIM; 308700; -.
DR MIM; 147950; -.
DR InterPro; IPR003961; FN_III.
DR InterPro; IPR002221; WAP.
DR Pfam; PF00041; fn3; 3.
DR Pfam; PF00095; wap; 1.
```

DR PRINTS; PR00003; 4DISULPHCORE.
DR SMART; SM00060; FN3; 4.
DR SMART; SM00217; WAP; 1.
DR PROSITE; PS00317; 4 DISULFIDE CORE; 1.
KW Cell adhesion; Glycoprotein; Serine protease inhibitor; Repeat;
KW Signal; Polymorphism; Disease mutation.
FT SIGNAL 1 20 POTENTIAL.
FT CHAIN 21 680 ANOSMIN 1.
FT DOMAIN 21 120 "CYSTEINE BOX".
FT DOMAIN 130 176 WAP.
FT DOMAIN 181 285 FIBRONECTIN TYPE-III 1.
FT DOMAIN 286 402 FIBRONECTIN TYPE-III 2.
FT DOMAIN 403 540 FIBRONECTIN TYPE-III 3.
FT DOMAIN 541 661 FIBRONECTIN TYPE-III 4.
FT CARBOHYD 71 71 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 209 209 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 300 300 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 470 470 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 553 553 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 564 564 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT VARIANT 267 267 N -> K (IN KAL1).
FT /FTID=VAR_007720.
FT E -> K (IN KAL1).
FT /FTID=VAR_012742.
FT I -> V.
FT /FTID=VAR_007721.
FT NN -> VR (IN REF. 4).
FT CONFLICT 70 71 E -> K (IN REF. 3).
FT CONFLICT 373 373 A -> R (IN REF. 3).
FT CONFLICT 540 540
FT SEQUENCE 680 AA; 76066 MW; 5D6ACC9F14B5F5F8 CRC64;

Query Match 28.3%; Score 92.5; DB 1; Length 680;
Best Local Similarity 34.6%; Pred. No. 0.0055;
Matches 18; Conservative 3; Mismatches 20; Indels 11; Gaps 2;

Qy 12 KPGSCPI-----ILIRCAMLNPNNRCLKDTDCPGIKKCEGSCGMACFVPQ 57
Db 130 KQGDPAPEKASGFAAACV-----ESCEVDNECSGVKKCCSNGCGHTCQVPK 176

RESULT 18
WAP_CAMDR STANDARD; PRT; 117 AA.
AC P09837;
DT 01-MAR-1989 (Rel. 10, Created)
DT 01-MAR-1989 (Rel. 10, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Whey acidic protein (WAP).
GN WAP.
OS Camelus dromedarius (Dromedary) (Arabian camel).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Tylopoda; Camelidae; Camelus.
OX NCBI_TaxID=9838;
RN [1]
RP SEQUENCE.
RC TISSUE=Milk;
RX MEDLINE=86300719; PubMed=3743571;
RA Beg O.U., von Bahr-Lindstrom H., Zaidi Z.H., Joernvall H.;
RT "A camel milk whey protein rich in half-cystine. Primary structure,
RT assessment of variations, internal repeat patterns, and relationships
RT with neurophysin and other active polypeptides.";
RL Eur. J. Biochem. 159:195-201(1986).
CC -!- FUNCTION: COULD BE A PROTEASE INHIBITOR.
CC -!- SUBCELLULAR LOCATION: Secreted.
CC -!- TISSUE SPECIFICITY: MILK-SPECIFIC; MAJOR PROTEIN COMPONENT OF MILK
CC WHEY.
CC -!- SIMILARITY: CONTAINS 2 WAP-TYPE DOMAINS.
DR PIR; A24178; A24178.
DR HSSP; O46655; 1CJH.
DR InterPro; IPR002221; WAP.
DR Pfam; PF00095; wap; 2.
DR PRINTS; PR00003; 4DISULPHCORE.
DR ProDom; PD001224; WAP; 1.

DR SMART; SM00217; WAP; 2.
DR PROSITE; PS00317; 4 DISULFIDE CORE; 2.
KW Milk; Whey; Protease inhibitor; Repeat; Phosphorylation.
FT DOMAIN 9 54 WAP 1.
FT DOMAIN 64 111 WAP 2.
FT DISULFID 13 41 BY SIMILARITY.
FT DISULFID 24 46 BY SIMILARITY.
FT DISULFID 28 40 BY SIMILARITY.
FT DISULFID 34 50 BY SIMILARITY.
FT DISULFID 68 99 BY SIMILARITY.
FT DISULFID 80 103 BY SIMILARITY.
FT DISULFID 86 98 BY SIMILARITY.
FT DISULFID 92 107 BY SIMILARITY.
SQ SEQUENCE 117 AA; 12564 MW; 2D9BB6A5A37A921B CRC64;

Query Match 27.7%; Score 90.5; DB 1; Length 117;
Best Local Similarity 37.5%; Pred. No. 0.0021;
Matches 21; Conservative 3; Mismatches 25; Indels 7; Gaps 2;

Qy 3 EPVKGPVSTKPGSCPIIL--IRCAMLNPNNRCLKDTDCPGIKKCEGSCGMACFVP 56
Db 60 EPV-----LKDGRCPPWQVOTPLTAKHCKLEKDCSRDDQCEGNKKCCFSSCAMRCLDP 110

RESULT 19
WFD2_CANFA STANDARD; PRT; 124 AA.
AC Q28894;
DT 15-JUL-1998 (Rel. 36, Created)
DT 15-JUL-1998 (Rel. 36, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Major epididymis-specific protein E4 precursor (CE4) (Epididymal
DE secretory protein E4) (WAP four-disulfide core domain protein 2).
GN WFD2.
OS Canis familiaris (Dog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Carnivora; Fissipedia; Canidae; Canis.
OX NCBI_TaxID=9615;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Epididymis;
RX MEDLINE=95263175; PubMed=7744511;
RA Ellerbrock K., Pera I., Hartung S., Ivell R.;
RT "Gene expression in the dog epididymis: a model for human epididymal
RT function.";
RL Int. J. Androl. 17:314-323(1994).
RN [2]
RP TISSUE SPECIFICITY.
RX MEDLINE=95263176; PubMed=7744512;
RA Pera I., Ivell R., Kirchhoff C.;
RT "Regional variation of specific gene expression in the dog epididymis
RT as revealed by in-situ transcript hybridization.";
RL Int. J. Androl. 17:324-330(1994).
CC -!- FUNCTION: POSSIBLE FUNCTION IN SPERM MATURATION.
CC -!- SUBCELLULAR LOCATION: Secreted (Potential).
CC -!- TISSUE SPECIFICITY: EPIDIDYMIS. HIGHEST LEVELS ARE FOUND
CC IN THE CAPUT AND PROXIMAL CAUDA REGIONS. LOWER LEVELS IN THE
CC DISTAL CAUDA. NOT DETECTED IN THE EFFERENT DUCTS.
CC -!- SIMILARITY: CONTAINS 2 WAP-TYPE DOMAINS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; S77395; AAB34264.1; -.
DR HSSP; O46655; 1CJH.
DR InterPro; IPR002221; WAP.
DR Pfam; PF00095; wap; 2.
DR PRINTS; PR00003; 4DISULPHCORE.

```
DR ProDom; PD001224; WAP; 1.
DR SMART; SM00217; WAP; 2.
DR PROSITE; PS00317; 4 DISULFIDE_CORE; 2.
KW Repeat; Signal; Glycoprotein.
FT SIGNAL 1 21 POTENTIAL.
FT CHAIN 22 124 MAJOR EPIDIDYMIS-SPECIFIC PROTEIN E4.
FT DOMAIN 32 74 WAP 1.
FT DOMAIN 76 124 WAP 2.
FT DISULFID 36 62 BY SIMILARITY.
FT DISULFID 45 66 BY SIMILARITY.
FT DISULFID 49 61 BY SIMILARITY.
FT DISULFID 55 70 BY SIMILARITY.
FT DISULFID 80 110 BY SIMILARITY.
FT DISULFID 93 114 BY SIMILARITY.
FT DISULFID 97 109 BY SIMILARITY.
FT DISULFID 103 119 BY SIMILARITY.
FT CARBOHYD 44 44 N-LINKED (GLCNAC... ) (POTENTIAL).
SQ SEQUENCE 124 AA; 12951 MW; 15AAF315BA13958C CRC64;

Query Match 26.9%; Score 88; DB 1; Length 124;
Best Local Similarity 38.3%; Pred. No. 0.0042;
Matches 18; Conservative 7; Mismatches 20; Indels 2; Gaps 2;

QY 12 KPGSCPILIRCAMLN-PNRCCLKDTDCPGIKKCCGSCG-MACFVP 56
| | | | | : | | | | | : | | | | | : | | | | |
Dd 76 KEGSCPQVNTDFPQLGLCQDQVDSHCPGLLKCCYNGCGKVCVTP 122

RESULT 20
EPPI_MOUSE STANDARD; PRT; 134 AA.
AC Q9DA01;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Eppin precursor (Epididymal protease inhibitor) (Serine protease
DE inhibitor-like with Kunitz and WAP domains 1).
GN SPINLW1.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=BALB/c; TISSUE=Testis, and Epididymis;
RA Sivashanmugam P., Hall S.H., Hamil K.G., French F.S., O'Rand M.G.,
RA Richardson R.T.;
RT "Characterization of monkey and mouse Eppin, a protease inhibitor from
RT epididymis and testis."
RL Submitted (FEB-2001) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Testis;
RX MEDLINE=21085660; PubMed=11217851;
RA Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
RA Arakawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,
RA Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamanaka I.,
RA Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R.,
RA Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,
RA Fleischmann W., Gaasterland T., Gissi C., King B., Kochiwa H.,
RA Kuehl P., Lewis S., Matsuo Y., Nikaido I., Pesole G., Quackenbush J.,
RA Schriml L.M., Staubli F., Suzuki R., Tomita M., Wagner L., Washio T.,
RA Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,
RA Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,
RA Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,
RA Gustincich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,
RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,
RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
RA Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,
RA Suzuki H., Toyo-oka K., Wang K.H., Weitz C., Whittaker C., Wilming L.,
RA Wynshaw-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kohtsuki S.,
RA Hayashizaki Y.;
RT "Functional annotation of a full-length mouse cDNA collection.";
```

```
RL Nature 409:685-690(2001).
CC -!- SUBCELLULAR LOCATION: Secreted (Potential).
CC -!- TISSUE SPECIFICITY: EXPRESSED IN EPIDIDYMIS AND TESTIS.
CC -!- SIMILARITY: CONTAINS 1 BPTI/KUNITZ INHIBITOR DOMAIN.
CC -!- SIMILARITY: CONTAINS 1 WAP-TYPE DOMAIN.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; AF346413; AAK31335.1; -.
CC EMBL; AK006296; BAB24514.1; -.
CC HSSP; P31713; 1SHP.
CC MGD; MGI:1922776; 1700024E17Rik.
CC InterPro; IPR002223; Kunitz_BPTI.
CC InterPro; IPR002221; WAP.
CC Pfam; PF00014; Kunitz_BPTI; 1.
CC Pfam; PF00095; wap; 1.
CC PRINTS; PR00759; BASICPTASE.
CC ProDom; PD000222; Kunitz_BPTI; 1.
CC SMART; SM00131; KU; 1.
CC SMART; SM00217; WAP; 1.
CC PROSITE; PS00317; 4 DISULFIDE_CORE; 1.
CC PROSITE; PS00280; BPTI_KUNITZ_1; 1.
CC PROSITE; PS00279; BPTI_KUNITZ_2; 1.
KW Serine protease inhibitor; Signal.
FT SIGNAL 1 21 POTENTIAL.
FT CHAIN 22 134 EPPIN.
FT DOMAIN 29 73 WAP.
FT DOMAIN 77 127 BPTI/KUNITZ INHIBITOR.
FT DISULFID 33 61 BY SIMILARITY.
FT DISULFID 40 65 BY SIMILARITY.
FT DISULFID 48 60 BY SIMILARITY.
FT DISULFID 54 69 BY SIMILARITY.
FT DISULFID 77 127 BY SIMILARITY.
FT DISULFID 86 110 BY SIMILARITY.
FT DISULFID 102 123 BY SIMILARITY.
SQ SEQUENCE 134 AA; 15470 MW; DFFEB63D4D4C427F CRC64;

Query Match 26.9%; Score 88; DB 1; Length 134;
Best Local Similarity 35.7%; Pred. No. 0.0045;
Matches 20; Conservative 5; Mismatches 27; Indels 4; Gaps 2;

QY 5 VKGPVSTK---PGSCPIILIRCAMLNPPNRCCLKDTDCPGIKKCCGSCGMACFVPQ 57
| | | | | : | | | | | : | | | | | : | | | | |
Dd 19 VQGPSLADLLFPRCPRFRECEH-QERDLCTDRDCPKKEKCCVFNCGKCLNPQ 73

RESULT 21
EPPI_HUMAN STANDARD; PRT; 133 AA.
AC Q95925; Q9HD30; Q96SD7;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Eppin precursor (Epididymal protease inhibitor) (Serine protease
DE inhibitor-like with Kunitz and WAP domains 1).
GN SPINLW1.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A. (ISOFORMS 1 AND 2).
RC TISSUE=Epididymis, and Testis;
RX MEDLINE=21297183; PubMed=11404006;
RA Richardson R.T., Sivashanmugam P., Hall S.H., Hamil K.G., Moore P.A.,
RA Ruben S.M., French F.S., O'Rand M.G.;
RT "Cloning and sequencing of human Eppin: a novel family of protease
```

RT inhibitors expressed in the epididymis and testis.";

RL Gene 270:93-102(2001).

RN [2]

RP SEQUENCE FROM N.A. (ISOFORM 1).

RA Stavrides G.S., Huckle E.J., Deloukas P.;

RL Submitted (SEP-1999) to the EMBL/GenBank/DBJ databases.

RN [3]

RP SEQUENCE FROM N.A.

RX MEDLINE=21638749; PubMed=11780052;

RA Deloukas P., Matthews L.H., Ashurst J., Burton J., Gilbert J.G.R., Jones M., Stavrides G., Almeida J.P., Babbage A.K., Bagguley C.L., Bailey J., Barlow K.F., Bates K.N., Beard L.M., Beare D.M., Beasley O.P., Bird C.P., Blakey S.E., Bridgeman A.M., Brown A.J., Buck D., Burrill W.D., Butler A.P., Carder C., Carter N.P., Chapman J.C., Clamp M., Clark G., Clark L.N., Clark S.Y., Clee C.M., Clegg S., Copley V.E., Collier R.E., Connor R.E., Corby N.R., Coulson A., Coville G.J., Deadman R., Dhani P.D., Dunn M., Ellington A.G., Frankland J.A., Fraser A., French L., Garner P., Grafham D.V., Griffiths C., Griffiths M.N.D., Gwilliam R., Hall R.E., Hammond S., Harley J.L., Heath P.D., Ho S., Holden J.L., Howden P.J., Huckle E., Hunt A.R., Hunt S.E., Jekosch K., Johnson C.M., Johnson D., Kay M.P., Kimberley A.M., King A., Knights A., Laird G.K., Lawlor S., Levasailho M.H., Leversha M.A., Lloyd C., Lloyd D.M., Lovell J.D., Marsh V.L., Martin S.L., McConnachie L.J., McLay K., McMurray A.A., Milne S.A., Mistry D., Moore M.J.F., Mullikin J.C., Nickerson T., Oliver K., Parker A., Patel R., Pearce T.A.V., Peck A.I., Phillimore B.J.C.T., Prathalingam S.R., Plumb R.W., Ramsay H., Rice C.M., Ross M.T., Scott C.E., Sehra H.K., Shownkeen R., Sims S., Skuce C.D., Smith M.L., Soderlund C., Steward C.A., Sulston J.E., Swann R.M., Sycamore N., Taylor R., Tee L., Thomas D.W., Thorpe A., Tracey A., Tromans A.C., Vaudin M., Wall M., Wallis J.M., Whitehead S.L., Whittaker P., Willey D.L., Williams L., Williams S.A., Wilming L., Wray P.W., Hubbard T., Durbin R.M., Bentley D.R., Beck S., Rogers J.;

RT "The DNA sequence and comparative analysis of human chromosome 20.";

RL Nature 414:865-871(2001).

CC -!- SUBCELLULAR LOCATION: Secreted (Potential).

CC -!- ALTERNATIVE PRODUCTS: 2 ISOFORMS; 1 (SHOWN HERE) AND 2; ARE PRODUCED BY ALTERNATIVE SPLICING.

CC -!- TISSUE SPECIFICITY: EXPRESSED IN EPIDIDYMIS AND TESTIS.

CC -!- SIMILARITY: CONTAINS 1 BPTI/KUNITZ INHIBITOR DOMAIN.

CC -!- SIMILARITY: CONTAINS 1 WAP-TYPE DOMAIN.

CC -----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

CC -----

CC EMBL; AF286370; AAG00548.1; -

DR EMBL; AF286369; AAG00547.1; -

DR EMBL; AF286368; AAG00546.1; -

DR EMBL; AL118493; CAB56343.1; -

DR EMBL; AL031663; CAB37635.1; -

DR EMBL; AL031663; CAC36265.1; -

DR HSSP; P00974; 1BPI.

DR Genew; HGNC:15932; SPINLW1.

DR InterPro; IPR002223; Kunitz_BPTI.

DR InterPro; IPR002221; WAP.

DR Pfam; PF00014; Kunitz_BPTI; 2.

DR Pfam; PF00095; wap; 2.

DR PRINTS; PR00759; BASICPTASE.

DR ProDom; PD000222; Kunitz_BPTI; 1.

DR SMART; SM00131; KU; 1.

DR SMART; SM00217; WAP; 1.

DR PROSITE; PS00317; 4 DISULFIDE CORE; 1.

DR PROSITE; PS00280; BPTI_KUNITZ_1; 1.

DR PROSITE; PS0279; BPTI_KUNITZ_2; 1.

KW Serine protease inhibitor; Signal; Alternative splicing.

FT SIGNAL 1 21 POTENTIAL.

FT CHAIN 22 133 EPPIN.

FT DOMAIN 29 73 WAP.

FT DOMAIN 77 127 BPTI/KUNITZ INHIBITOR.

FT DISULFID 33 61 BY SIMILARITY.

FT DISULFID 40 65 BY SIMILARITY.

FT DISULFID 48 60 BY SIMILARITY.

FT DISULFID 54 69 BY SIMILARITY.

FT DISULFID 77 127 BY SIMILARITY.

FT DISULFID 86 110 BY SIMILARITY.

FT DISULFID 102 123 BY SIMILARITY.

FT VARSPLIC 1 31 MGSSGLLSLLVFLVLLANVQGPGLTDWLFPR -> MLSKAH

FT GCKTALSIG (IN ISOFORM 2).

SQ SEQUENCE 133 AA; 15284 MW; F7831B203366D9DC CRC64;

Query Match 26.0%; Score 85; DB 1; Length 133;

Best Local Similarity 40.4%; Pred. No. 0.0094;

Matches 21; Conservative 2; Mismatches 25; Indels 4; Gaps 2;

QY 5 VKGPVSTK--PGSCPILIRCAMLPPNRCCLKOTDCPGIKKCCGSCGMAC 53

Db 19 VQGPGLTDWLFPRRCPKIRECE-FQERDVCTKDRQCQDNKKCCVFCGKKC 69

RESULT 22

EPPI_MACMU

ID EPPI_MACMU STANDARD; PRT; 133 AA.

AC Q9BDL1;

DT 15-JUN-2002 (Rel. 41, Created)

DT 15-JUN-2002 (Rel. 41, Last sequence update)

DT 15-JUN-2002 (Rel. 41, Last annotation update)

DE Eppin precursor (Epididymal protease inhibitor) (Serine protease inhibitor-like with Kunitz and WAP domains 1).

GN SPINLW1.

OS Macaca mulatta (Rhesus macaque).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Cercopitheidae; Cercopithecinae; Macaca.

OC NCBI_TaxID=9544;

OX [1]

RN SEQUENCE FROM N.A.

RC TISSUE=Testis, and Epididymis;

RA Sivashanmugam P., Hall S.H., Hamil K.G., French F.S., O'Rand M.G., Richardson R.T.;

RA "Characterization of monkey and mouse Eppin, a protease inhibitor from epididymis and testis.";

RT Submitted (FEB-2001) to the EMBL/GenBank/DBJ databases.

RL -!- SUBCELLULAR LOCATION: Secreted (Potential).

CC -!- TISSUE SPECIFICITY: EXPRESSED IN EPIDIDYMIS AND TESTIS.

CC -!- SIMILARITY: CONTAINS 1 BPTI/KUNITZ INHIBITOR DOMAIN.

CC -!- SIMILARITY: CONTAINS 1 WAP-TYPE DOMAIN.

CC -----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

CC -----

CC EMBL; AF346414; AAK31336.1; -

DR HSSP; P00974; 1BPI.

DR InterPro; IPR002223; Kunitz_BPTI.

DR InterPro; IPR002221; WAP.

DR Pfam; PF00014; Kunitz_BPTI; 1.

DR Pfam; PF00095; wap; 1.

DR PRINTS; PR00759; BASICPTASE.

DR ProDom; PD000222; Kunitz_BPTI; 1.

DR SMART; SM00131; KU; 1.

DR SMART; SM00217; WAP; 1.

DR PROSITE; PS00280; BPTI_KUNITZ_1; 1.

DR PROSITE; PS0279; BPTI_KUNITZ_2; 1.

KW Serine protease inhibitor; Signal.

FT SIGNAL 1 21 POTENTIAL.

FT CHAIN 22 133 EPPIN.


```

FT  CONFLICT 132 132 MISSING (IN REF. 2).
SQ  SEQUENCE 211 AA; 23079 MW; 605E7E6B0FC64FE5 CRC64;

Query Match 24.6%; Score 80.5; DB 1; Length 211;
Best Local Similarity 30.4%; Pred. No. 0.042;
Matches 14; Conservative 6; Mismatches 15; Indels 11; Gaps 1;

QY 8 PVSTKPGSCPILLIRCAMLNPPNRLKDTDCPGIKKCCGSGCMAC 53
   | : |||:|
Db 61 PRTLPPGAC-----QATRCQADSECPRHRRCCYNGCAYAC 95

RESULT 25
WAP_MOUSE
ID  WAP_MOUSE STANDARD; PRT; 134 AA.
AC  P01173; P70230; Q61023;
DT  21-JUL-1986 (Rel. 01, Created)
DT  16-OCT-2001 (Rel. 40, Last sequence update)
DT  16-OCT-2001 (Rel. 40, Last annotation update)
DE  Whey acidic protein precursor (WAP).
GN  WAP.
OS  Mus musculus (Mouse).
OC  Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC  Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX  NCBI_TaxID=10090;
RN  [1]
RP  SEQUENCE FROM N.A.
RX  MEDLINE=85062841; PubMed=6095207;
RA  Campbell S.M., Rosen J.M., Hennighausen L.G., Strech-Jurk U.,
RA  Sippel A.E.;
RT  "Comparison of the whey acidic protein genes of the rat and mouse.";
RL  Nucleic Acids Res. 12:8685-8697(1984).
RN  [2]
RP  SEQUENCE FROM N.A.
RX  MEDLINE=82196900; PubMed=6896234;
RA  Hennighausen L.G., Sippel A.E.;
RT  "Mouse whey acidic protein is a novel member of the family of 'four-
RT  disulfide core' proteins.";
RL  Nucleic Acids Res. 10:2677-2684(1982).
RN  [3]
RP  SEQUENCE FROM N.A.
RC  STRAIN=GR;
RA  Hennighausen L.;
RL  Submitted (OCT-1995) to the EMBL/GenBank/DBJ databases.
RN  [4]
RP  CHARACTERIZATION.
RC  STRAIN=YBR;
RX  MEDLINE=82052974; PubMed=6975276;
RA  Piletz J.E., Heinlen M., Ganschow R.E.;
RT  "Biochemical characterization of a novel whey protein from murine
RT  milk.";
RL  J. Biol. Chem. 256:11509-11516(1981).
CC  -!- FUNCTION: COULD BE A PROTEASE INHIBITOR. MAY PLAY AN IMPORTANT
CC  ROLE IN MAMMARY GLAND DEVELOPMENT AND TISSUE REMODELING.
CC  -!- SUBCELLULAR LOCATION: Secreted.
CC  -!- TISSUE SPECIFICITY: MILK-SPECIFIC; MAJOR PROTEIN COMPONENT OF MILK
CC  WHEY.
CC  -!- PTM: NO PHOSPHATE OR CARBOHYDRATE BINDING COULD BE DETECTED;
CC  HOWEVER, BOTH CHOLESTEROL AND TRIGLYCERIDE ARE ASSOCIATED WITH THE
CC  MOUSE PROTEIN.
CC  -!- SIMILARITY: CONTAINS 2 WAP-TYPE DOMAINS.
CC  -----
CC  This SWISS-PROT entry is copyright. It is produced through a collaboration
CC  between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC  the European Bioinformatics Institute. There are no restrictions on its
CC  use by non-profit institutions as long as its content is in no way
CC  modified and this statement is not removed. Usage by and for commercial
CC  entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC  or send an email to license@isb-sib.ch).
CC  -----
DR  EMBL; X01157; CAA25604.1; -.
DR  EMBL; X01158; CAA25604.1; JOINED.
DR  EMBL; X01159; CAA25604.1; JOINED.

```

21-JUL-1986 (Rel. 01, Created)
21-JUL-1986 (Rel. 01, Last sequence update)
16-OCT-2001 (Rel. 40, Last annotation update)
Chelonianin (Basic protease inhibitor) (RTPI).
Caretta caretta (Loggerhead).
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Testudines; Cryptodira; Chelonioidae; Cheloniidae; Caretta.
NCBI_TaxID=8467;
[1]
SEQUENCE.
TISSUE=Egg white;
Kato I., Tominaga N.;
"Trypsin-subtilisin inhibitor from red sea turtle eggwhite consists
of two tandem domains -- one Kunitz -- one of a new family.";
Fed. Proc. 38:832-832(1979).
-!- FUNCTION: THE FIRST DOMAIN INHIBITS TRYPSIN; THE SECOND ONE
INHIBITS SUBTILISIN.
-!- SIMILARITY: CONTAINS 1 BPTI/KUNITZ INHIBITOR DOMAIN.
-!- SIMILARITY: CONTAINS 1 WAP-TYPE DOMAIN.
-!- CAUTION: AS THE PAPER ONLY INDICATES THE SPECIES AS "RED SEA
TURTLE", THE SPECIES INDICATED HERE IS THEREFORE AN INFERENCE.
PIR; A01224; TITTOR.
HSSP; P00974; 6PTI.
InterPro; IPR002223; Kunitz_BPTI.
InterPro; IPR002221; WAP.

DR Pfam; PF00014; Kunitz_BPTI; 1.
DR Pfam; PF00095; wap; 1.
DR PRINTS; PR00003; 4DISULPHCORE.
DR PRINTS; PR00759; BASICPTASE.
DR ProDom; PD000222; Kunitz_BPTI; 1.
DR SMART; SM00131; KU; 1.
DR SMART; SM00217; WAP; 1.
DR PROSITE; PS00317; 4 DISULFIDE CORE; 1.
DR PROSITE; PS00280; BPTI_KUNITZ_1; 1.
DR PROSITE; PS00279; BPTI_KUNITZ_2; 1.
KW Serine protease inhibitor.
FT DOMAIN 8 58 BPTI/KUNITZ INHIBITOR.
FT DOMAIN 63 105 WAP.
FT MOD_RES 1 1 PYRROLIDONE CARBOXYLIC ACID.
FT DISULFID 8 58 BY SIMILARITY.
FT DISULFID 17 41 BY SIMILARITY.
FT DISULFID 33 54 BY SIMILARITY.
FT DISULFID 67 92 BY SIMILARITY.
FT DISULFID 76 97 BY SIMILARITY.
FT DISULFID 80 93 BY SIMILARITY.
FT DISULFID 86 101 BY SIMILARITY.
FT ACT_SITE 18 19 REACTIVE BOND (TRYPSIN).
SQ SEQUENCE 110 AA; 11916 MW; 269436243813418E CRC64;

Query Match 22.8%; Score 74.5; DB 1; Length 110;
Best Local Similarity 34.8%; Pred. No. 0.11;
Matches 16; Conservative 4; Mismatches 15; Indels 11; Gaps 2;

QY 12 KPSCPIILIRCAMLPPNRL-----KDTDCPGIKKCEGSCGMAC 53
Db 63 RGVCP-----KTSVPGICLHGCDSDCKEKGKCFDGCYIC 101

RESULT 29
ITR1_MOMCO
ID ITR1_MOMCO STANDARD; PRT; 34 AA.
AC P82408;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Trypsin inhibitor I (MCOTI-I).
OS Momordica cochinchinensis (Spiny bitter cucumber).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eurosids I; Cucurbitales; Cucurbitaceae; Momordica.
OX NCBI_TaxID=3674;
RN [1]
RP SEQUENCE, AND MASS SPECTROMETRY.
RC TISSUE=Seed;
RX MEDLINE=20263490; PubMed=10801322;
RA Hernandez J.-F., Gagnon J., Chiche L., Nguyen T.M., Andrieu J.-P.,
Heitz A., Trinh T., Pham T.T.C., Le Nguyen D.;
RT "Squash trypsin inhibitors from Momordica cochinchinensis exhibit an
RT atypical macrocyclic structure."
RL Biochemistry 39:5722-5730(2000).
CC -!- FUNCTION: INHIBITS TRYPSIN; PROBABLY PARTICIPATES IN A PLANT
CC DEFENSE MECHANISM.
CC -!- PTM: THIS IS A CYCLIC PEPTIDE.
CC -!- MASS SPECTROMETRY: MW=3480.7; MW_ERR=0.3; METHOD=Electrospray.
CC -!- SIMILARITY: BELONGS TO THE SQUASH FAMILY OF SERINE PROTEASE
CC INHIBITORS.
DR HSP; P12071; 2ETI.
DR InterPro; IPR000737; Squash.
DR Pfam; PF00299; squash; 1.
DR PRINTS; PR00293; SQUASHINHBTR.
DR ProDom; PD003401; Squash; 1.
DR SMART; SM00286; PTI; 1.
DR PROSITE; PS00286; SQUASH INHIBITOR; 1.
KW Serine protease inhibitor.
FT ACT_SITE 10 11 REACTIVE BOND.
FT DISULFID 8 25 BY SIMILARITY.
FT DISULFID 15 27 BY SIMILARITY.
FT DISULFID 21 33 BY SIMILARITY.

SQ SEQUENCE 34 AA; 3505 MW; 8F7D0B4C162C935A CRC64;

Query Match 20.5%; Score 67; DB 1; Length 34;
Best Local Similarity 42.1%; Pred. No. 0.28;
Matches 16; Conservative 2; Mismatches 10; Indels 10; Gaps 2;

QY 14 GSCPIILIRCAMLPPNRLKDTDCPGIKK-CEGSCG 50
Db 6 GVC PKIL-----QRCRRSDCPGACICRNGYCG 34

RESULT 30
FBN2_HUMAN
ID FBN2_HUMAN STANDARD; PRT; 2911 AA.
AC P35556;
DT 01-JUN-1994 (Rel. 29, Created)
DT 01-FEB-1996 (Rel. 33, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Fibrillin 2 precursor.
GN FBN2.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Euthera; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=94165150; PubMed=8120105;
RA Zhang H., Apfelroth S.D., Hu W., Davis E.C., Sanguinetti C.,
Bonadio J., Mecham R.P., Ramirez F.;
RT "Structure and expression of fibrillin-2, a novel microfibrillar
RT component preferentially located in elastic matrices."
RL J. Cell Biol. 124:855-863(1994).
RN [2]
RP SEQUENCE OF 752-1505 FROM N.A.
RX MEDLINE=91304567; PubMed=1852206;
RA Lee B., Godfrey M., Vitale E., Hori H., Mattei M.-G., Sarfarazi M.,
Tsipouras P., Ramirez F., Hollister D.;
RT "Linkage of Marfan syndrome and a phenotypically related disorder to
RT two different fibrillin genes."
RL Nature 352:330-334(1991).
RN [3]
RP VARIANTS CCA TYR-1252 AND SER-1433, AND VARIANT ILE-964.
RX MEDLINE=96083599; PubMed=7493032;
RA Putnam E.A., Zhang H., Ramirez F., Milewicz D.M.;
RT "Fibrillin-2 (FBN2) mutations result in the Marfan-like disorder,
RT congenital contractual arachnodactyly."
RL Nat. Genet. 11:456-458(1995).
RN [4]
RP VARIANTS CCA HIS-1114.
RX MEDLINE=98407789; PubMed=9737771;
RA Babcock D., Gasner C., Francke U., Maslen C.;
RT "A single mutation that results in an asp-to-his substitution and
RT partial exon skipping in a family with congenital contractual
RT arachnodactyly."
RL Hum. Genet. 103:22-28(1998).
RN [5]
RP VARIANTS CCA PHE-1141 AND TRP-1252.
RX MEDLINE=20259236; PubMed=10797416;
RA Belleh S., Zhou G., Wang M., Der Kaloustian V.M., Pagon R.A.,
Godfrey M.;
RT "Two novel fibrillin-2 mutations in congenital contractual
RT arachnodactyly."
RL Am. J. Med. Genet. 92:7-12(2000).
CC -!- FUNCTION: STRUCTURAL COMPONENT OF CONNECTIVE TISSUE MICROFIBRILS
CC THAT BINDS CALCIUM. FIBRILLIN-2-CONTAINING MICROFIBRILS REGULATE
CC THE EARLY PROCESS OF ELASTIC FIBER ASSEMBLY.
CC -!- DISEASE: DEFECTS IN FBN2 ARE THE CAUSE OF CONGENITAL CONTRACTURAL
CC ARACHNOACTYLY (CCA) (ALSO KNOWN AS BEALS SYNDROME). CCA IS
CC PHENOTYPICALLY SIMILAR TO MARFAN SYNDROME, BUT DOES NOT EFFECT THE
CC AORTA AND THE EYES.
CC -!- SIMILARITY: CONTAINS 47 EGF-LIKE DOMAINS.
CC -!- SIMILARITY: CONTAINS 7 TGF-BETA BINDING PROTEIN DOMAINS.


```
Db 2018 EVALPGSCSPGTQCNLEGSFRC--ICPPGYEVKSENCDINECEDPNICLFGSCTNTP 2075
QY 50 -GMACFVP 56
Db 2076 GGFQCLCP 2083

RESULT 31
BAR3_CHITE
ID BAR3_CHITE STANDARD; PRT; 1700 AA.
AC Q03376;
DT 01-OCT-1993 (Rel. 27, Created)
DT 01-OCT-1993 (Rel. 27, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Balbiani ring protein 3 precursor.
GN BR3.
OS Chironomus tentans (Midge).
OC Eukaryota; Metazoa; Arthropoda; Mandibulata; Pancrustacea; Hexapoda;
OC Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Nematocera;
OC Chironomoidea; Chironomidae; Chironominae; Chironomus.
OX NCBI_TaxID=7153;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Salivary gland;
RX MEDLINE=90172404; PubMed=1689777;
RA Paulsson G., Lendahl U., Galli J., Ericsson C., Wieslander L.;
RT "The Balbiani ring 3 gene in Chironomus tentans has a diverged
RL J. Mol. Biol. 211:331-349(1990).
CC -!- FUNCTION: USED BY THE LARVAE TO CONSTRUCT A SUPRAMOLECULAR
CC STRUCTURE, THE LARVAL TUBE. BALBIANI RING PROTEIN 3 COULD PLAY A
CC ROLE AS A TRANSPORT PROTEIN THAT BINDS TO OTHER PROTEINS
CC INTRACELLULARLY AND IN THE GLAND LUMEN IN ORDER TO PREVENT THESE
CC FROM FORMING WATER-INSOLUBLE FIBERS TOO EARLY.
CC -!- SUBCELLULAR LOCATION: Secreted.
CC -!- TISSUE SPECIFICITY: SALIVARY GLAND.
CC -!- DOMAIN: HAS 82 APPROXIMATE REPEATS OF CYS-X-CYS-X-CYS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; X52263; CAA36506.1; -
DR PIR; S08167; S08167.
DR HSSP; P15358; 1SKZ.
DR InterPro; IPR004153; CXCXC_repeat.
DR Pfam; PF03128; CXCXC; 71.
KW Repeat; Signal.
FT SIGNAL 1 20 POTENTIAL.
FT CHAIN 21 1700 BALBIANI RING PROTEIN 3.
SQ SEQUENCE 1700 AA; 186145 MW; 34202B28521B0815 CRC64;

Query Match 19.9%; Score 65; DB 1; Length 1700;
Best Local Similarity 32.7%; Pred. No. 11;
Matches 17; Conservative 7; Mismatches 14; Indels 14; Gaps 4;

QY 13 PGSCPIILIRCAMLNPPN--RCLKDT---DCPGIKKC-----CEGSCGMAC 53
Db 267 PNTCE---CGCAQLNCPDNKKANKETCQCECKEVKCKNGGQVFCCKDSCSVC 315

RESULT 32
ITR3_MOMCO
ID ITR3_MOMCO STANDARD; PRT; 30 AA.
AC P82410;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DE Balbiani ring protein 3 precursor.
GN BR3.
OS Chironomus tentans (Midge).
OC Eukaryota; Metazoa; Arthropoda; Mandibulata; Pancrustacea; Hexapoda;
OC Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Nematocera;
OC Chironomoidea; Chironomidae; Chironominae; Chironomus.
OX NCBI_TaxID=7153;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Salivary gland;
RX MEDLINE=90172404; PubMed=1689777;
RA Paulsson G., Lendahl U., Galli J., Ericsson C., Wieslander L.;
RT "The Balbiani ring 3 gene in Chironomus tentans has a diverged
RL J. Mol. Biol. 211:331-349(1990).
CC -!- FUNCTION: USED BY THE LARVAE TO CONSTRUCT A SUPRAMOLECULAR
CC STRUCTURE, THE LARVAL TUBE. BALBIANI RING PROTEIN 3 COULD PLAY A
CC ROLE AS A TRANSPORT PROTEIN THAT BINDS TO OTHER PROTEINS
CC INTRACELLULARLY AND IN THE GLAND LUMEN IN ORDER TO PREVENT THESE
CC FROM FORMING WATER-INSOLUBLE FIBERS TOO EARLY.
CC -!- SUBCELLULAR LOCATION: Secreted.
CC -!- TISSUE SPECIFICITY: SALIVARY GLAND.
CC -!- DOMAIN: HAS 82 APPROXIMATE REPEATS OF CYS-X-CYS-X-CYS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; X52263; CAA36506.1; -
DR PIR; S08167; S08167.
DR HSSP; P15358; 1SKZ.
DR InterPro; IPR004153; CXCXC_repeat.
DR Pfam; PF03128; CXCXC; 71.
KW Repeat; Signal.
FT SIGNAL 1 20 POTENTIAL.
FT CHAIN 21 1700 BALBIANI RING PROTEIN 3.
SQ SEQUENCE 1700 AA; 186145 MW; 34202B28521B0815 CRC64;

Query Match 19.9%; Score 65; DB 1; Length 1700;
Best Local Similarity 32.7%; Pred. No. 11;
Matches 17; Conservative 7; Mismatches 14; Indels 14; Gaps 4;
```

```
DE Trypsin inhibitor III (MCOTI-III).
OS Momordica cochinchinensis (Spiny bitter cucumber).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eurosids I; Cucurbitales; Cucurbitaceae; Momordica.
OX NCBI_TaxID=3674;
RN [1]
RP SEQUENCE, AND MASS SPECTROMETRY.
RC TISSUE=Seed;
RX MEDLINE=20263490; PubMed=10801322;
RA Hernandez J.-F., Gagnon J., Chiche L., Nguyen T.M., Andrieu J.-P.,
RA Heitz A., Trinh T., Pham T.T.C., Le Nguyen D.;
RT "Squash trypsin inhibitors from Momordica cochinchinensis exhibit an
RT atypical macrocyclic structure.";
RL Biochemistry 39:5722-5730(2000).
CC -!- FUNCTION: INHIBITS TRYPSIN; PROBABLY PARTICIPATES IN A PLANT
CC DEFENSE MECHANISM.
CC -!- MASS SPECTROMETRY: MW=3379.6; MW ERR=0.5; METHOD=Electrospray.
CC -!- SIMILARITY: BELONGS TO THE SQUASH FAMILY OF SERINE PROTEASE
CC INHIBITORS.
DR HSSP; P01074; IBXJ.
DR InterPro; IPR000737; Squash.
DR Pfam; PF00299; squash; 1.
DR PRINTS; PR00293; SQUASHINHBTR.
DR ProDom; PD003401; Squash; 1.
DR SMART; SM00286; PTI; 1.
DR PROSITE; PS00286; SQUASH_INHIBITOR; 1.
KW Serine protease inhibitor.
FT MOD_RES 1 1 PYRROLIDONE CARBOXYLIC ACID.
FT ACT_SITE 6 7 REACTIVE BOND.
FT DISULFID 4 21 BY SIMILARITY.
FT DISULFID 11 23 BY SIMILARITY.
FT DISULFID 17 29 BY SIMILARITY.
SQ SEQUENCE 30 AA; 3403 MW; D5E13476B1A9AA6E CRC64;

Query Match 19.6%; Score 64; DB 1; Length 30;
Best Local Similarity 40.5%; Pred. No. 0.53;
Matches 15; Conservative 4; Mismatches 8; Indels 10; Gaps 2;

QY 15 SCPIILIRCAMLNPPNRLKDTDCPGIKKCCG-GSCG 50
Db 3 ACPRIL-----KKCRDSDCPGECICKENGYCG 30

RESULT 33
ITR2_MOMCO
ID ITR2_MOMCO STANDARD; PRT; 34 AA.
AC P82409;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Trypsin inhibitor II (MCOTI-II).
OS Momordica cochinchinensis (Spiny bitter cucumber).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eurosids I; Cucurbitales; Cucurbitaceae; Momordica.
OX NCBI_TaxID=3674;
RN [1]
RP SEQUENCE, AND MASS SPECTROMETRY.
RC TISSUE=Seed;
RX MEDLINE=20263490; PubMed=10801322;
RA Hernandez J.-F., Gagnon J., Chiche L., Nguyen T.M., Andrieu J.-P.,
RA Heitz A., Trinh T., Pham T.T.C., Le Nguyen D.;
RT "Squash trypsin inhibitors from Momordica cochinchinensis exhibit an
RT atypical macrocyclic structure.";
RL Biochemistry 39:5722-5730(2000).
CC -!- FUNCTION: INHIBITS TRYPSIN; PROBABLY PARTICIPATES IN A PLANT
CC DEFENSE MECHANISM.
CC -!- PTM: A CYCLIC SUCCINIMIDE MAY BE FORMED BETWEEN ASP-4 AND GLY-5
CC RESIDUES WHICH CAN BE TRANSFORMED TO A BETA-ASPARTYL BOND. SO
CC THERE EXISTS THREE ISOFORMS OF MCOTI-II, NORMAL, WITH SUCCINIMIDE
CC OR WITH A BETA-ASPARTYL BOND.
CC -!- PTM: THIS IS A CYCLIC PEPTIDE.
```

```
CC -!- MASS SPECTROMETRY: MW=3453; MW_ERR=0.2; METHOD=Electrospray.
CC -!- SIMILARITY: BELONGS TO THE SQUASH FAMILY OF SERINE PROTEASE
CC INHIBITORS.
DR HSSP; P01074; 1BXJ.
DR InterPro; IPR000737; Squash.
DR Pfam; PF00299; squash; 1.
DR PRINTS; PR00293; SQUASHINHBTR.
DR ProDom; PD003401; Squash; 1.
DR SMART; SM00286; PPI; 1.
DR PROSITE; PS00286; SQUASH_INHIBITOR; 1.
KW Serine protease inhibitor.
FT ACT SITE 10 11 REACTIVE BOND.
FT DISULFID 8 25 BY SIMILARITY.
FT DISULFID 15 27 BY SIMILARITY.
FT DISULFID 21 33 BY SIMILARITY.
SQ SEQUENCE 34 AA; 3477 MW; 8F7D0B4C048BB93A CRC64;

Query Match 19.6%; Score 64; DB 1; Length 34;
Best Local Similarity 39.5%; Pred. No. 0.58;
Matches 15; Conservative 3; Mismatches 10; Indels 10; Gaps 2;

QY 14 GSCPIILIRCAMLNPPNRLKDTDCPGIKK-CEGSCG 50
   ||||| :|||:|||||
Db 6 GVCPKIL-----KKCRSDSDCPGACICRNGVCG 34

RESULT 34
FBN2_MOUSE STANDARD; PRT; 2907 AA.
AC Q61555; Q63957;
DT 15-DEC-1998 (Rel. 37, Created)
DT 15-DEC-1998 (Rel. 37, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Fibrillin 2 precursor.
GN FBN2 OR FBN-2.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Euthera; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=95263670; PubMed=7744963;
RA Li X., Pereira L., Zhang H., Sanguinetti C., Ramirez F., Bonadio J.,
RA Francke U.;
RT "Fibrillin genes map to regions of conserved mouse/human synteny on
RT mouse chromosomes 2 and 18.";
RL Genomics 18:667-672(1993).
CC -!- FUNCTION: STRUCTURAL COMPONENT OF CONNECTIVE TISSUE MICROFIBRILS
CC THAT BINDS CALCIUM. FIBRILLIN-1-CONTAINING MICROFIBRILS PROVIDE
CC LONG-TERM FORCE BEARING STRUCTURAL SUPPORT.
CC -!- SIMILARITY: CONTAINS 47 EGF-LIKE DOMAINS.
CC -!- SIMILARITY: CONTAINS 7 TGF-BETA BINDING PROTEIN DOMAINS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; L39790; AAA74908.1; -
CC EMBL; S69359; AAC60685.1; -
CC HSSP; P35555; 1EMN.
CC MGD; MGI:95490; Fbn2.
DR InterPro; IPR000152; Asx_hydroxyl.
```

```
DR InterPro; IPR000561; EGF-like.
DR InterPro; IPR001881; EGF Ca.
DR InterPro; IPR001438; EGF_II.
DR InterPro; IPR002212; Fibril-assoc.
DR Pfam; PF00008; EGF; 46.
DR Pfam; PF00683; TB; 9.
DR PRINTS; PR00010; EGFLOOD.
DR SMART; SM00179; EGF_CA; 43.
DR SMART; SM00001; EGF_like; 3.
DR PROSITE; PS00010; ASX_HYDROXYL; 43.
DR PROSITE; PS00022; EGF_1; 2.
DR PROSITE; PS01186; EGF_2; 36.
DR PROSITE; PS01187; EGF_CA; 43.
KW Extracellular matrix; Calcium-binding; Glycoprotein; EGF-like domain;
KW Repeat; Signal; Multigene family.
FT SIGNAL 1 28 POTENTIAL.
FT CHAIN 29 2907 FIBRILLIN 2.
FT DOMAIN 111 142 EGF-LIKE 1.
FT DOMAIN 145 176 EGF-LIKE 2.
FT DOMAIN 176 208 EGF-LIKE 3.
FT DOMAIN 276 317 EGF-LIKE 4, CALCIUM-BINDING.
FT DOMAIN 318 359 EGF-LIKE 5, CALCIUM-BINDING.
FT REPEAT 360 426 TGFBP 1.
FT DOMAIN 487 527 EGF-LIKE 6.
FT DOMAIN 528 567 EGF-LIKE 7, CALCIUM-BINDING.
FT DOMAIN 568 609 EGF-LIKE 8, CALCIUM-BINDING.
FT DOMAIN 610 650 EGF-LIKE 9, CALCIUM-BINDING.
FT DOMAIN 651 691 EGF-LIKE 10, CALCIUM-BINDING.
FT REPEAT 692 760 TGFBP 2.
FT DOMAIN 761 802 EGF-LIKE 11, CALCIUM-BINDING.
FT DOMAIN 803 844 EGF-LIKE 12, CALCIUM-BINDING.
FT DOMAIN 845 883 EGF-LIKE 13, CALCIUM-BINDING.
FT DOMAIN 948 989 EGF-LIKE 14, CALCIUM-BINDING.
FT REPEAT 990 1065 TGFBP 3.
FT DOMAIN 1066 1107 EGF-LIKE 15, CALCIUM-BINDING.
FT DOMAIN 1108 1150 EGF-LIKE 16, CALCIUM-BINDING.
FT DOMAIN 1151 1192 EGF-LIKE 17, CALCIUM-BINDING.
FT DOMAIN 1193 1234 EGF-LIKE 18, CALCIUM-BINDING.
FT DOMAIN 1235 1275 EGF-LIKE 19, CALCIUM-BINDING.
FT DOMAIN 1276 1317 EGF-LIKE 20, CALCIUM-BINDING.
FT DOMAIN 1318 1359 EGF-LIKE 21, CALCIUM-BINDING.
FT DOMAIN 1360 1400 EGF-LIKE 22, CALCIUM-BINDING.
FT DOMAIN 1401 1441 EGF-LIKE 23, CALCIUM-BINDING.
FT DOMAIN 1442 1483 EGF-LIKE 24, CALCIUM-BINDING.
FT DOMAIN 1484 1524 EGF-LIKE 25, CALCIUM-BINDING.
FT DOMAIN 1525 1565 EGF-LIKE 26, CALCIUM-BINDING.
FT REPEAT 1566 1642 TGFBP 4.
FT DOMAIN 1643 1684 EGF-LIKE 27, CALCIUM-BINDING.
FT DOMAIN 1685 1726 EGF-LIKE 28, CALCIUM-BINDING.
FT REPEAT 1727 1800 TGFBP 5.
FT DOMAIN 1801 1842 EGF-LIKE 29, CALCIUM-BINDING.
FT DOMAIN 1843 1884 EGF-LIKE 30, CALCIUM-BINDING.
FT DOMAIN 1885 1926 EGF-LIKE 31, CALCIUM-BINDING.
FT DOMAIN 1927 1965 EGF-LIKE 32, CALCIUM-BINDING.
FT DOMAIN 1966 2008 EGF-LIKE 33, CALCIUM-BINDING.
FT DOMAIN 2009 2048 EGF-LIKE 34, CALCIUM-BINDING.
FT DOMAIN 2049 2090 EGF-LIKE 35, CALCIUM-BINDING.
FT REPEAT 2091 2163 TGFBP 6.
FT DOMAIN 2164 2205 EGF-LIKE 36, CALCIUM-BINDING.
FT DOMAIN 2206 2245 EGF-LIKE 37, CALCIUM-BINDING.
FT DOMAIN 2246 2286 EGF-LIKE 38, CALCIUM-BINDING.
FT DOMAIN 2287 2330 EGF-LIKE 39, CALCIUM-BINDING.
FT DOMAIN 2331 2372 EGF-LIKE 40, CALCIUM-BINDING.
FT REPEAT 2373 2441 TGFBP 7.
FT DOMAIN 2442 2483 EGF-LIKE 41, CALCIUM-BINDING.
FT DOMAIN 2484 2524 EGF-LIKE 42, CALCIUM-BINDING.
FT DOMAIN 2525 2563 EGF-LIKE 43, CALCIUM-BINDING.
FT DOMAIN 2564 2606 EGF-LIKE 44, CALCIUM-BINDING.
FT DOMAIN 2607 2646 EGF-LIKE 45, CALCIUM-BINDING.
FT DOMAIN 2647 2687 EGF-LIKE 46, CALCIUM-BINDING.
FT DOMAIN 2688 2727 EGF-LIKE 47, CALCIUM-BINDING.
FT DISULFID 115 124 BY SIMILARITY.
FT DISULFID 119 130 BY SIMILARITY.
```

SQ SEQUENCE 202 AA; 23258 MW; 644113CE7AEC2AC2 CRC64;

DR InterPro; IPR001438; EGF II.
DR InterPro; IPR000800; NotCh.
DR Pfam; PF00008; EGF; 27.
DR Pfam; PF00023; ank; 6.
DR Pfam; PF00066; notch; 2.
DR PRINTS; PR01415; ANKYRIN.
DR PRINTS; PR00010; EGFBLD.
DR PRINTS; PR01452; NOTCH.
DR SMART; SM00248; ANK; 5.
DR SMART; SM00179; EGF_CA; 11.
DR SMART; SM00001; EGF_like; 15.
DR SMART; SM00004; NL; 2.
DR PROSITE; PS50088; ANK_REPEAT; 5.
DR PROSITE; PS50297; ANK_REPEAT; 1.
DR PROSITE; PS00010; ASX_HYDROXYL; 11.
DR PROSITE; PS00022; EGF_1; 28.
DR PROSITE; PS01186; EGF_2; 21.
DR PROSITE; PS01187; EGF_CA; 9.
KW Receptor; Transcription regulation; Activator; Differentiation;
KW Developmental protein; Repeat; ANK repeat; EGF-like domain;
KW Transmembrane; Glycoprotein; Signal; Phosphorylation; Proto-oncogene.
FT SIGNAL 1 20
FT CHAIN 21 1964 NEUROGENIC LOCUS NOTCH HOMOLOG PROTEIN 4.
FT CHAIN 1411 1964 TRANSFORMING PROTEIN INT-3.
FT CHAIN 1428 1964 NOTCH EXTRACELLULAR TRUNCATION.
FT CHAIN 1463 1964 NOTCH INTRACELLULAR DOMAIN.
FT DOMAIN 21 1443 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 1444 1464 POTENTIAL.
FT DOMAIN 1465 1964 CYTOPLASMIC (POTENTIAL).
FT DOMAIN 21 60 EGF-LIKE 1.
FT DOMAIN 61 112 EGF-LIKE 2.
FT DOMAIN 115 152 EGF-LIKE 3.
FT DOMAIN 153 189 EGF-LIKE 4.
FT DOMAIN 191 229 EGF-LIKE 5.
FT DOMAIN 231 271 EGF-LIKE 6.
FT DOMAIN 273 309 EGF-LIKE 7.
FT DOMAIN 311 350 EGF-LIKE 8.
FT DOMAIN 352 388 EGF-LIKE 9.
FT DOMAIN 389 427 EGF-LIKE 10.
FT DOMAIN 429 470 EGF-LIKE 11.
FT DOMAIN 472 508 EGF-LIKE 12.
FT DOMAIN 510 546 EGF-LIKE 13.
FT DOMAIN 548 584 EGF-LIKE 14.
FT DOMAIN 586 622 EGF-LIKE 15.
FT DOMAIN 623 656 EGF-LIKE 16.
FT DOMAIN 658 686 EGF-LIKE 17.
FT DOMAIN 688 724 EGF-LIKE 18.
FT DOMAIN 726 762 EGF-LIKE 19.
FT DOMAIN 764 800 EGF-LIKE 20.
FT DOMAIN 803 839 EGF-LIKE 21.
FT DOMAIN 841 877 EGF-LIKE 22.
FT DOMAIN 878 924 EGF-LIKE 23.
FT DOMAIN 926 962 EGF-LIKE 24.
FT DOMAIN 964 1000 EGF-LIKE 25.
FT DOMAIN 1002 1040 EGF-LIKE 26.
FT DOMAIN 1042 1081 EGF-LIKE 27.
FT DOMAIN 1083 1122 EGF-LIKE 28.
FT DOMAIN 1126 1167 EGF-LIKE 29.
FT REPEAT 1168 1208 LIN/NOTCH 1.
FT REPEAT 1209 1242 LIN/NOTCH 2.
FT REPEAT 1243 1282 LIN/NOTCH 3.
FT REPEAT 1283 1324 ANK 1.
FT REPEAT 1325 1365 ANK 2.
FT REPEAT 1366 1406 ANK 3.
FT REPEAT 1407 1447 ANK 4.

Query Match 19.1%; Score 62.5; DB 1; Length 1964;
Best Local Similarity 29.9%; Pred. No. 24;
Matches 20; Conservative 7; Mismatches 11; Indels 29; Gaps 4;

QY 10 STKPGSCPIILIRCAMLN-----PPNRCLKD---TDCPGIKKCCGSC 49
Db 1101 SPKPGSPPL-----CACLSGFGGPDCLTPAPPCCGPPSPCLHNGTCTETPGL-----GNP 1151

QY 50 GMACFVP 56
Db 1152 GFQCTCP 1158
RESULT 37
MT2A MYTED
ID MT2A MYTED STANDARD; PRT; 71 AA.
AC P80253; Q9GU31;
DT 01-FEB-1994 (Rel. 28, Created)
DT 01-FEB-1994 (Rel. 28, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Metallothionein 20-III isoform A (MT-20-III) (MT-I).
OS Mytilus edulis (Blue mussel), and
OS Mytilus galloprovincialis (Mediterranean mussel).
OC Eukaryota; Metazoa; Mollusca; Bivalvia; Pteriomorpha; Mytiloidea;
OC Mytiloidea; Mytilidae; Mytilus.
OX NCBI_TaxID=6550, 29158;
RN [1]
RP SEQUENCE.
RC SPECIES=M.edulis;
RX MEDLINE=94062828; PubMed=8243463;
RA Mackay E.A., Overnell J., Dunbar B., Davidson I., Hunziker P.E.,
RA Kaegi J.H.R., Fothergill J.E.;
RT "Complete amino acid sequences of five dimeric and four monomeric
forms of metallothionein from the edible mussel Mytilus edulis.";
RL Eur. J. Biochem. 218:183-194 (1993).
RN [2]
RP SEQUENCE FROM N.A.
RC SPECIES=M.galloprovincialis;
RA Ceratto N., Dondero F., Burlando B., Van de Loo J.W.H.P.,
RA Viarengo A.;
RT "Metallothionein genomic DNA of the mussel Mytilus
galloprovincialis.";
RL Submitted (OCT-1999) to the EMBL/GenBank/DBJ databases.
CC -!- FUNCTION: THE METALLOTHIONEINS ARE INVOLVED IN THE CELLULAR
CC SEQUESTRATION OF TOXIC METAL IONS.
CC -!- SUBUNIT: HOMODIMER.
CC -!- INDUCTION: BY CADMIUM.
CC -!- SIMILARITY: BELONGS TO THE METALLOTHIONEIN SUPERFAMILY; FAMILY 2.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See [http://www.isb-sib.ch/](http://www.isb-sib.ch/announce/)
CC or send an email to license@isb-sib.ch).

DR EMBL; AF199020; AAC28538.1; -.
DR PIR; S39422; S39422.
DR InterPro; IPR003019; Metallthion.
DR InterPro; IPR001008; Mtlthion_misc.
DR Pfam; PF00131; metalthio; 1.
DR PRINTS; PR00875; MTMOLLUSC.
KW Metal-binding; Metal-thiolate cluster; Cadmium.
FT INIT_MET 0 0 BY SIMILARITY.
SQ SEQUENCE 71 AA; 6899 MW; D4BFD9E599857657 CRC64;

Query Match 19.0%; Score 62; DB 1; Length 71;
Best Local Similarity 39.0%; Pred. No. 1.8;
Matches 16; Conservative 2; Mismatches 11; Indels 12; Gaps 3;

QY 13 PGSCPIILIRCAMLNPPNRCLKDTDCPGIKKCCGSCGMAC 53
Db 1 PGPCNCI-----ETNVCICGTGCSG--KCCQ--CGDAC 29

RESULT 38
MT2B MYTED
ID MT2B MYTED STANDARD; PRT; 71 AA.
AC P80258;

DR Pfam; PF00041; fn3; 3.
DR Pfam; PF00069; pkinase; 1.
DR Pfam; PF00757; Furin-like; 1.
DR Pfam; PF01030; Recep_L domain; 2.
DR PRINTS; PR00014; FNTYPEIII.
DR PRINTS; PR00109; TYRKINASE.
DR ProDom; PD000001; Euk_pkinase; 1.
DR SMART; SM00060; FN3; 3.
DR SMART; SM00261; FU; 1.
DR SMART; SM00219; TyrKC; 1.
DR PROSITE; PS00107; PROTEIN KINASE ATP; 1.
DR PROSITE; PS00109; PROTEIN KINASE_TYR; 1.
DR PROSITE; PS00239; RECEPTOR_TYR_KIN_II; 1.
DR PROSITE; PS50011; PROTEIN KINASE_DOM; 1.
KW Transferrase; Tyrosine-protein kinase; Receptor; Transmembrane;
KW Glycoprotein; ATP-binding; Phosphorylation; Signal.
FT SIGNAL 1 29
FT CHAIN 30 716
FT INSULIN-LIKE PEPTIDE RECEPTOR, ALPHA-
FT CHAIN (POTENTIAL).
FT PROPEP 717 720
FT CHAIN 721 1363
FT INSULIN-LIKE PEPTIDE RECEPTOR, BETA-
FT CHAIN (POTENTIAL).
FT DOMAIN 721 928
FT TRANSMEM 929 949
FT DOMAIN 950 1363
FT DOMAIN 994 1283
FT NP_BIND 1000 1008
FT BINDING 1028 1028
FT ACT_SITE 1148 1148
FT MOD_RES 1174 1174
FT CARBOHYD 51 51
FT CARBOHYD 97 97
FT CARBOHYD 137 137
FT CARBOHYD 278 278
FT CARBOHYD 483 483
FT CARBOHYD 599 599
FT CARBOHYD 617 617
FT CARBOHYD 665 665
FT CARBOHYD 666 666
FT CARBOHYD 711 711
FT CARBOHYD 732 732
FT CARBOHYD 736 736
FT CARBOHYD 743 743
FT CARBOHYD 816 816
FT CARBOHYD 885 885
FT CARBOHYD 898 898
SQ SEQUENCE 1363 AA; 154104 MW; 238120B4EAB1ED65 CRC64;
Query Match 18.8%; Score 61.5; DB 1; Length 1363;
Best Local Similarity 23.7%; Pred. No. 23;
Matches 18; Conservative 7; Mismatches 8; Indels 43; Gaps 5;
QY 14 GSCPIILIRCAMLNPPN-----RCLKDTCFPGI-----KKC----- 44
Db 252 GEC---LIQC-----PPDTYQYKDRRCITEECPTNSVWKLHHRKCIPCCPSGYTTDIN 304
QY 45 -----CEGSCGMAC 53
Db 305 NPLRLTECEGQCPKSC 320
RESULT 41
TFF2_HUMAN
ID TFF2_HUMAN STANDARD; PRT; 129 AA.
AC Q03403; Q15854;
DT 01-OCT-1994 (Rel. 30, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Trefoil factor 2 precursor (Spasmolytic polypeptide) (SP).
GN TFF2 OR SML1.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Stomach;
RX MEDLINE=90151615; PubMed=2303034;
RA Tomasetto C., Rio M.C., Gautier C., Wolf C., Hareuveni M., Chambon P.,
RA Lathe R.;
RT "hSP, the domain-duplicated homolog of pS2 protein, is co-expressed
RT with pS2 in stomach but not in breast carcinoma.";
RL EMBO J. 9:407-414(1990).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=97224512; PubMed=9070946;
RA Seib T., Blin N., Hilgert K., Seifert M., Theisinger B., Engel M.,
RA Dooley S., Zang K.D., Welter C.;
RT "The three human trefoil genes TFF1, TFF2, and TFF3 are located
RT within a region of 55 kb on chromosome 21q22.3.";
RL Genomics 40:200-202(1997).
RN [3]
RP SEQUENCE FROM N.A.
RX MEDLINE=20408883; PubMed=10950923;
RA Berry A., Scott H.S., Kudoh J., Talior I., Korostishevsky M.,
RA Wattenhofer M., Guipponi M., Barras C., Rossier C., Shibuya K.,
RA Wang J., Kawasaki K., Asakawa S., Minoshima S., Shimizu N.,
RA Antonarakis S., Bonne-Tamir B.;
RT "Refined localization of autosomal recessive nonsyndromic deafness
RT DFNB10 locus using 34 novel microsatellite markers, genomic structure,
RT and exclusion of six known genes in the region.";
RL Genomics 68:22-29(2000).
CC -!- FUNCTION: INHIBITS GASTROINTESTINAL MOTILITY AND GASTRIC ACID
CC SECRETION. COULD FUNCTION AS A STRUCTURAL COMPONENT OF GASTRIC
CC MUCUS, POSSIBLY BY STABILIZING GLYCOPROTEINS IN THE MUCUS GEL
CC THROUGH INTERACTIONS WITH CARBOHYDRATE SIDE CHAINS (BY
CC SIMILARITY).
CC -!- SUBCELLULAR LOCATION: Secreted.
CC -!- TISSUE SPECIFICITY: STOMACH.
CC -!- SIMILARITY: CONTAINS 2 P-TYPE (TREFOIL) DOMAINS.
CC -!- DATABASE: NAME=Atlas Genet. Cytogenet. Oncol. Haematol.;
CC WWW="http://www.infobiogen.fr/services/chromancer/Genes/TFF2ID264.html".

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).

CC EMBL; X51698; CAA35995.1; --
DR EMBL; U47292; AAB05397.1; --
DR EMBL; U47289; AAB05397.1; JOINED.
DR EMBL; U47290; AAB05397.1; JOINED.
DR EMBL; U47291; AAB05397.1; JOINED.
DR EMBL; AB038162; BAB13730.1; --
DR PIR; S12371; S12371.
DR HSSP; P01359; 1PCP.
DR Genew; HGNC:11756; TFF2.
DR MIM; 182590; --
DR InterPro; IPR000519; P_trefoil.
DR Pfam; PF00088; trefoil_2.
DR PRINTS; PR00680; PTREFOIL.
DR SMART; SM00018; P; 2.
DR PROSITE; PS00025; P_TREFOIL; 2.
KW Repeat; Signal.
FT SIGNAL 1 23
FT CHAIN 24 129
FT DOMAIN 30 72
FT DOMAIN 80 121
FT DISULFID 29 127
FT DISULFID 31 58
FT DISULFID 42 57
FT DISULFID 52 69
FT DISULFID 81 107
FT POTENTIAL.
FT TREFOIL FACTOR 2.
FT P-TYPE 1.
FT P-TYPE 2.
FT BY SIMILARITY.
FT BY SIMILARITY.
FT BY SIMILARITY.
FT BY SIMILARITY.

```
FT DISULFID 91 106 BY SIMILARITY.
FT DISULFID 101 118 BY SIMILARITY.
FT CONFLICT 1 19 MGRDQAALLVGLCA -> RHGTARRPAPGSAPRPGA
FT CONFLICT 122 122 K -> N (IN REF. 1).
FT CONFLICT 129 AA; 14284 MW; 6A296A11AD9CCC5D CRC64;
SQ SEQUENCE 129 AA; 14284 MW; 6A296A11AD9CCC5D CRC64;

Query Match 18.7%; Score 61; DB 1; Length 129;
Best Local Similarity 36.7%; Pred. No. 3.7;
Matches 18; Conservative 6; Mismatches 9; Indels 16; Gaps 5;

QY 10 STKPGSCPIILIRCAMLPNRLKDTDC--PGL--KKCEGSCGMACF 54
| | | | | : | | | | | : | | | | | : | | | | |
Db 23 SEKSPSC-----QCSRLSPHNR-----TNCGFPGITSDQCFDNGC---CF 59
| | | | | : | | | | | : | | | | | : | | | | |

RESULT 42
YMW2 CAEEL STANDARD; PRT; 1246 AA.
AC P34504; P34505; P34506; P90907;
DT 01-FEB-1994 (Rel. 28, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Hypothetical protein K04H4.2 in chromosome III.
GN K04H4.2.
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC Rhabditidae; Peloderinae; Caenorhabditis.
OX NCBI_TaxID=6239;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Bristol N2;
RX MEDLINE=94150718; PubMed=7906398;
RA Wilson R., Ainscough R., Anderson K., Baynes C., Berks M., Coulson A.,
RA Bonfield J., Burton J., Connell M., Copsey T., Cooper J., Fraser A.,
RA Craxton M., Dear S., Du Z., Durbin R., Favello A., Fraser A.,
RA Fulton L., Gardner A., Green P., Hawkins T., Hillier L., Jier M.,
RA Johnston L., Jones M., Kershaw J., Kirsten J., Laister N.,
RA Latreille P., Lightning J., Lloyd C., Mortimore B., O'Callaghan M.,
RA Parsons J., Percy C., Rifken L., Roopra A., Saunders D., Shownkeen R.,
RA Sims M., Smaldon N., Smith A., Smith M., Sonhammer E., Staden R.,
RA Sulston J., Thierry-Mieg J., Thomas K., Vaudin M., Vaughan K.,
RA Waterston R., Watson A., Weinstock L., Wilkinson-Sproat J.,
RA Wohldman P.;
RA "2.2 Mb of contiguous nucleotide sequence from chromosome III of C.
RT elegans.";
RL Nature 368:32-38 (1994).
RN [2]
RP REVISIONS.
RA Durbin R.;
RL Submitted (NOV-1998) to the EMBL/GenBank/DBJ databases.
CC -!- ALTERNATIVE PRODUCTS: 2 isoforms; a (shown here) and b; may be
CC produced by alternative splicing.
CC -!- SIMILARITY: TO CHITIN-BINDING MOTIFS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; Z27078; CAA81587.1; -.
CC EMBL; Z27078; CAA81588.1; -.
CC PIR; S40992; S40992.
CC PIR; S40993; S40993.
CC PIR; S40994; S40994.
CC HSSP; P10968; 2CWG.
CC WormPep; K04H4.2a; CE19967.
CC WormPep; K04H4.2b; CE19968.
CC InterPro; IPR002557; Chitin_bind_PeA.
CC InterPro; IPR002899; WR1/EB.
```

```
DR SMART; SM00494; ChtBD2; 1.
DR SMART; SM00289; WR1; 14.
KW Hypothetical protein; Alternative splicing.
FT VARSPLIC 166 529
ITPTPTRTTEPPKRLCLSSNTDEVNLSGASSSSATCGGTN
ANCTDEDCPTTFKCYGCGCKLAVCPRLTAVKFTCKTQYH
CRANEHCFGGCCPKTIELAVIKSQVLTMSKDNHETKETEK
LIIGDCEVDTRVKKCDIDIICPEMSECVDGICCKQPPKARC
GNGLMALSIHVHCSLSDDCPIASRCCEYKCCPFLSESADST
SDSVGETTPVIIKEEIIISTATKVMKKVDKTSVGSINKKCL
STQRCDLHTLCPPDFTCSLSGCKCKLNIHCPDGTVPETSCQ
SASNDHCPSSSHKCTLLNKEHFACCVSPGLVVEGSVTAEV
SSECPICSVVEVDPRFGTSCRYSLQCFSPYFCNQRGQ -> Q
(IN ISOFORM B).
FT VARSPLIC 570 715
DSDTNTSTNPSQPETTTKNTKSNSSKKHRKPKKKDVPD
LSDPLLQNDFFPIGPGYGFPEHLNLDLVLIRAOGDGECPA
GLHCDTAINLCCPLLLPLTDPKPKKRRKTKERRKQKQDGNEM
EASANFDSDDPARFSSYSCGCMG -> VG (IN
ISOFORM B).
SQ SEQUENCE 1246 AA; 130610 MW; 4FA1A17D3F9606C4 CRC64;

Query Match 18.7%; Score 61; DB 1; Length 1246;
Best Local Similarity 47.6%; Pred. No. 24;
Matches 10; Conservative 2; Mismatches 9; Indels 0; Gaps 0;

QY 32 CLKDTDCPGIKKCCGSCGMA 52
| | | | | : | | | | | : | | | | | : | | | | |
Db 209 CTSDEDCPTTFKCYGCGCKLA 229
| | | | | : | | | | | : | | | | | : | | | | |

RESULT 43
ITR1 MOMCH STANDARD; PRT; 30 AA.
AC P10294;
DT 01-MAR-1989 (Rel. 10, Created)
DT 01-MAR-1989 (Rel. 10, Last sequence update)
DT 01-JUN-1994 (Rel. 29, Last annotation update)
DE Trypsin inhibitor I (MCTI-I).
OS Momordica charantia (Bitter melon) (Bitter melon).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eurosids I; Cucurbitales; Cucurbitaceae; Momordica.
OX NCBI_TaxID=3673;
RN [1]
RP SEQUENCE.
RC TISSUE=Seed; PubMed=2738047;
RX MEDLINE=89291812; PubMed=2738047;
RA Hara S., Makino J., Ikenaka T.;
RT "Amino acid sequences and disulfide bridges of serine proteinase
RT inhibitors from bitter melon (Momordica charantia LINN.) seeds.";
RL J. Biochem. 105:88-92 (1989).
CC -!- SIMILARITY: BELONGS TO THE SQUASH FAMILY OF SERINE PROTEASE
CC INHIBITORS.
CC PIR; JX0057; JX0057.
CC HSSP; P01074; 1BXJ.
CC InterPro; IPR000737; Squash.
CC Pfam; PF00299; squash; 1.
CC PRINTS; PR00293; SQUASHINHBTR.
CC ProDom; PD003401; Squash; 1.
CC SMART; SM00286; PTI; 1.
CC PROSITE; PS00286; SQUASH INHIBITOR; 1.
KW Serine protease inhibitor.
FT ACT SITE 6 7 REACTIVE BOND.
FT DISULFID 4 21 BY SIMILARITY.
FT DISULFID 11 23 BY SIMILARITY.
FT DISULFID 17 29 BY SIMILARITY.
SQ SEQUENCE 30 AA; 3413 MW; D5E1344365E34E9A CRC64;

Query Match 18.3%; Score 60; DB 1; Length 30;
Best Local Similarity 38.9%; Pred. No. 1.4;
Matches 14; Conservative 3; Mismatches 9; Indels 10; Gaps 2;

QY 16 CPIILIRCAMLPNRLKDTDCPGIKK-CEGSCG 50
| | | | | : | | | | | : | | | | | : | | | | |
```


DR PROSITE; PS50060; MAM_2; 3.
vw signal. Glycoprotein: Transmembrane: EGF-like domain; Cell adhesion;

RESULT 47
MT EISFO
ID MT EISFO STANDARD; PRT; 75 AA.
AC P81695; O96822;
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Cadmium-metallothionein (MT) (Fragment).
GN MT.
OS Eisenia foetida (Common brandling worm) (Common dung-worm).
OC Eukaryota; Metazoa; Annelida; Clitellata; Oligochaeta; Haplotax
OC Lumbricina; Lumbricidae; Eisenia.
OX NCBI_TaxID=6396;
RN [1]
RP SEQUENCE OF 1-41.
RA Gruber C., Dallinger R., Hunziker P., Gehrig P., Berger B.,
RA Stuerzenbaum S.;
RT "(Cd)-metallothionein from Eisenia foetida: evidence for
RT posttranslational processing to a functional one-domain protein
RL Submitted (FEB-1999) to the SWISS-PROT data bank.
RN [2]
RP SEQUENCE OF 2-75 FROM N.A.


```
FT DISULFID 280 291 BY SIMILARITY.
FT DISULFID 285 300 BY SIMILARITY.
FT DISULFID 302 311 BY SIMILARITY.
FT DISULFID 318 332 BY SIMILARITY.
FT DISULFID 326 341 BY SIMILARITY.
FT DISULFID 343 352 BY SIMILARITY.

Query Match 18.2%; Score 59.5; DB 1; Length 2003;
Best Local Similarity 23.4%; Pred. No. 51;
Matches 22; Conservative 5; Mismatches 16; Indels 51; Gaps 4;

Qy 10 STKPGSCPIILIRCAMLN-----PPNRCLKDTDC----- 38
Db 1105 SPKPGFPP-----RCACLSGYGGPDCLTPAPKGGPPSPCLYNGSCSETTGLGGPGFRCS 1160

Qy 39 -----PGIKKC-----CEGSCGMACFVP 56
Db 1161 CPHSSPGRCQKPGAKGCEGRSGDGACDAGCSGP 1194

RESULT 50
FBN1_MOUSE
ID FBN1_MOUSE STANDARD; PRT; 2871 AA.
AC Q61554; Q60826;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Fibrillin 1 precursor.
GN FBN1 OR FBN-1.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=95130561; PubMed=7829516;
RA Yin W., Germiller J., Sanguinetti C., Smiley E., Pangilinan T.,
RA Pereira L., Ramirez F., Bonadio J.;
RT "Primary structure and developmental expression of Fbn-1, the mouse
RL J. Biol. Chem. 270:1798-1806(1995).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=CD-1; TISSUE=Kidney;
RA Ota K., Kumar A., Wada J., Liu Z., Kanwar Y.S.;
RL Submitted (APR-1995) to the EMBL/GenBank/DBJ databases.
CC -!- FUNCTION: STRUCTURAL COMPONENT OF CONNECTIVE TISSUE MICROFIBRILS
CC THAT BINDS CALCIUM. FIBRILLIN-1-CONTAINING MICROFIBRILS PROVIDE
CC LONG-TERM FORCE BEARING STRUCTURAL SUPPORT.
CC -!- PTM: FORMS INTERMOLECULAR DISULFIDE BONDS EITHER WITH OTHER
CC FIBRILLIN-1 MOLECULES OR WITH OTHER COMPONENTS OF THE
CC MICROFIBRILS (BY SIMILARITY).
CC -!- SIMILARITY: CONTAINS 47 EGF-LIKE DOMAINS.
CC -!- SIMILARITY: CONTAINS 7 TGF-BETA BINDING PROTEIN DOMAINS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; L29454; AAA56840.1; -
CC EMBL; U22493; AAA64217.1; -
CC HSSP; P35555; 1APJ.
CC MGD; MGI:95489; Fbn1.
CC InterPro; IPR000152; Asx_hydroxyl.
CC InterPro; IPR000561; EGF-like.
CC InterPro; IPR001881; EGF_Ca.
CC InterPro; IPR002212; Fibril-assoc.
CC Pfam; PF00008; EGF; 46.
CC Pfam; PF00683; TB; 9.
```

```
DR SMART; SM00179; EGF_CA; 42.
DR SMART; SM00001; EGF_Like; 4.
DR PROSITE; PS00010; ASX_HYDROXYL; 43.
DR PROSITE; PS00022; EGF_1; 2.
DR PROSITE; PS01186; EGF_2; 38.
DR PROSITE; PS01187; EGF_CA; 43.
KW Extracellular matrix; Calcium-binding; Glycoprotein; EGF-like domain;
KW Repeat; Signal; Multigene family.
FT SIGNAL 1 27 POTENTIAL.
FT CHAIN 28 2871 FIBRILLIN 1.
FT DOMAIN 81 112 EGF-LIKE 1.
FT DOMAIN 115 146 EGF-LIKE 2.
FT DOMAIN 147 178 EGF-LIKE 3.
FT DOMAIN 246 287 EGF-LIKE 4.
FT DOMAIN 288 329 EGF-LIKE 5.
FT DOMAIN 330 401 TGFBP 1.
FT DOMAIN 402 446 PRO-RICH.
FT DOMAIN 449 489 EGF-LIKE 6.
FT DOMAIN 490 529 EGF-LIKE 7.
FT DOMAIN 530 571 EGF-LIKE 8.
FT DOMAIN 572 612 EGF-LIKE 9.
FT DOMAIN 613 653 EGF-LIKE 10.
FT DOMAIN 656 721 TGFBP 2.
FT DOMAIN 723 764 EGF-LIKE 11.
FT DOMAIN 765 806 EGF-LIKE 12.
FT DOMAIN 807 846 EGF-LIKE 13.
FT DOMAIN 910 951 EGF-LIKE 14.
FT DOMAIN 952 1018 TGFBP 3.
FT DOMAIN 1028 1069 EGF-LIKE 15.
FT DOMAIN 1070 1112 EGF-LIKE 16.
FT DOMAIN 1113 1154 EGF-LIKE 17.
FT DOMAIN 1155 1196 EGF-LIKE 18.
FT DOMAIN 1197 1237 EGF-LIKE 19.
FT DOMAIN 1238 1279 EGF-LIKE 20.
FT DOMAIN 1280 1321 EGF-LIKE 21.
FT DOMAIN 1322 1362 EGF-LIKE 22.
FT DOMAIN 1363 1403 EGF-LIKE 23.
FT DOMAIN 1404 1445 EGF-LIKE 24.
FT DOMAIN 1446 1486 EGF-LIKE 25.
FT DOMAIN 1487 1527 EGF-LIKE 26.
FT DOMAIN 1528 1599 TGFBP 4.
FT DOMAIN 1606 1647 EGF-LIKE 27.
FT DOMAIN 1648 1688 EGF-LIKE 28.
FT DOMAIN 1689 1758 TGFBP 5.
FT DOMAIN 1766 1807 EGF-LIKE 29.
FT DOMAIN 1808 1848 EGF-LIKE 30.
FT DOMAIN 1849 1890 EGF-LIKE 31.
FT DOMAIN 1891 1929 EGF-LIKE 32.
FT DOMAIN 1930 1972 EGF-LIKE 33.
FT DOMAIN 1973 2012 EGF-LIKE 34.
FT DOMAIN 2013 2054 EGF-LIKE 35.
FT DOMAIN 2055 2121 TGFBP 6.
FT DOMAIN 2127 2165 EGF-LIKE 36.
FT DOMAIN 2166 2205 EGF-LIKE 37.
FT DOMAIN 2206 2246 EGF-LIKE 38.
FT DOMAIN 2247 2290 EGF-LIKE 39.
FT DOMAIN 2291 2332 EGF-LIKE 40.
FT DOMAIN 2333 2400 TGFBP 7.
FT DOMAIN 2402 2443 EGF-LIKE 41.
FT DOMAIN 2444 2484 EGF-LIKE 42.
FT DOMAIN 2485 2523 EGF-LIKE 43.
FT DOMAIN 2524 2566 EGF-LIKE 44.
FT DOMAIN 2567 2606 EGF-LIKE 45.
FT DOMAIN 2607 2647 EGF-LIKE 46.
FT DOMAIN 2648 2687 EGF-LIKE 47.
FT DISULFID 85 94 BY SIMILARITY.
FT DISULFID 89 100 BY SIMILARITY.
FT DISULFID 102 111 BY SIMILARITY.
FT DISULFID 119 129 BY SIMILARITY.
FT DISULFID 123 134 BY SIMILARITY.
FT DISULFID 136 145 BY SIMILARITY.
FT DISULFID 150 160 BY SIMILARITY.
FT DISULFID 154 166 BY SIMILARITY.
```


RA van Vlijmen I.M., van Haren B., Miller W.L., Bristow J.;
RT "A recessive form of the Ehlers-Danlos syndrome caused by tenascin-X
RL deficiency.";
CC New Engl. J. Med. 345:1167-1175(2001).
CC -!- FUNCTION: APPEARS TO MEDIATE INTERACTIONS BETWEEN CELLS AND THE
CC EXTRACELLULAR MATRIX. SUBSTRATE-ADHESION MOLECULE THAT APPEARS TO
CC INHIBIT CELL MIGRATION. MAY PLAY A ROLE IN SUPPORTING THE GROWTH
CC OF EPITHELIAL TUMORS.
CC -!- SUBCELLULAR LOCATION: Extracellular matrix.
CC -!- ALTERNATIVE PRODUCTS: AT LEAST 2 ISOFORMS; XB (SHOWN HERE) AND
CC XB-SHORT; ARE PRODUCED BY ALTERNATIVE SPLICING.
CC -!- TISSUE SPECIFICITY: HIGHLY EXPRESSED IN FETAL ADRENAL, IN FETAL
CC TESTIS, FETAL SMOOTH, STRIATED AND CARDIAC MUSCLE. XB-SHORT IS
CC ONLY EXPRESSED IN THE ADRENAL GLAND.
CC -!- DISEASE: ASSOCIATION WITH CONGENITAL ADRENAL HYPERPLASIA.
CC -!- DISEASE: Defects in TNXB are the cause of an autosomal recessive
CC form of Ehlers-Danlos syndrome (EDS) manifesting with
CC hyperextensible skin, hypermobile joints, and tissue fragility as
CC observed in classical EDS, but lacking atrophic scars and delayed
CC wound healing.
CC -!- SIMILARITY: CONTAINS 19 EGF-LIKE DOMAINS.
CC -!- SIMILARITY: CONTAINS 32 FIBRONECTIN TYPE III-LIKE DOMAINS.
CC -!- SIMILARITY: CONTAINS 1 FIBRINOGEN C-TERMINAL DOMAIN.
CC -!- CAUTION: THERE ARE TWO GENES FOR TN-X: TNXA AND TNXB. TNXA IS A
CC PARTIAL GENE WHICH CAN SOMETIMES RECOMBINE WITH TNXB.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; U89337; AAB47488.1; -;
DR EMBL; AF019413; AAB67981.1; -;
DR EMBL; X71923; CAA50739.1; -;
DR EMBL; Y13782; CAA74109.1; -;
DR EMBL; Y13783; CAA74110.1; -;
DR EMBL; U24488; AAB41287.1; -;
DR EMBL; U52696; AAC50889.1; -;
DR EMBL; M25813; AAA35884.1; -;
DR HSSP; P02671; 1FZD.
DR Genew; HGNC:11976; TNXB.
DR MIM; 600985; -;
DR MIM; 606408; -;
DR InterPro; IPR000561; EGF-like.
DR InterPro; IPR003961; FN III.
DR InterPro; IPR002181; Fibrinogen C.
DR InterPro; IPR002049; Laminin_EGF.
DR Pfam; PF00008; EGF; 15.
DR Pfam; PF00041; fn3; 33.
DR Pfam; PF00147; fibrinogen C; 1.
DR PRINTS; PR00011; EGFLAMININ.
DR SMART; SM00181; EGF; 8.
DR SMART; SM00001; EGF_like; 1.
DR SMART; SM00186; FBG; 1.
DR SMART; SM00060; FN3; 31.
DR PROSITE; PS00022; EGF_1; 18.
DR PROSITE; PS01186; EGF_2; 19.
DR PROSITE; PS00514; FIBRIN AG_C DOMAIN; 1.
KW Glycoprotein; Cell adhesion; Repeat; EGF-like domain; Coiled coil;
KW Extracellular matrix; Alternative splicing; Signal;
KW Ehlers-Danlos syndrome.
FT SIGNAL 1 23 POTENTIAL.
FT CHAIN 24 4289 TENASCIN-X.
FT DOMAIN 156 168 EGF-LIKE 1 (INCOMPLETE).
FT DOMAIN 183 213 EGF-LIKE 2.
FT DOMAIN 214 244 EGF-LIKE 3.
FT DOMAIN 245 275 EGF-LIKE 4.
FT DOMAIN 276 306 EGF-LIKE 5.
FT DOMAIN 307 337 EGF-LIKE 6.
FT DOMAIN 338 368 EGF-LIKE 7.

FT DOMAIN 369 400 EGF-LIKE 8.
FT DOMAIN 400 430 EGF-LIKE 9.
FT DOMAIN 431 461 EGF-LIKE 10.
FT DOMAIN 462 492 EGF-LIKE 11.
FT DOMAIN 493 523 EGF-LIKE 12.
FT DOMAIN 524 554 EGF-LIKE 13.
FT DOMAIN 555 585 EGF-LIKE 14.
FT DOMAIN 586 616 EGF-LIKE 15.
FT DOMAIN 617 647 EGF-LIKE 16.
FT DOMAIN 648 679 EGF-LIKE 17.
FT DOMAIN 684 714 EGF-LIKE 18.
FT DOMAIN 715 746 EGF-LIKE 19.
FT DOMAIN 792 872 FIBRONECTIN TYPE-III 1.
FT DOMAIN 901 922 COILED COIL (POTENTIAL).
FT DOMAIN 941 1021 FIBRONECTIN TYPE-III 2.
FT DOMAIN 1047 1127 FIBRONECTIN TYPE-III 3.
FT DOMAIN 1149 1226 FIBRONECTIN TYPE-III 4.
FT DOMAIN 1246 1327 FIBRONECTIN TYPE-III 5.
FT DOMAIN 1348 1429 FIBRONECTIN TYPE-III 6.
FT DOMAIN 1459 1540 FIBRONECTIN TYPE-III 7.
FT DOMAIN 1561 1642 FIBRONECTIN TYPE-III 8.
FT DOMAIN 1659 1736 FIBRONECTIN TYPE-III 9.
FT DOMAIN 1756 1836 FIBRONECTIN TYPE-III 10.
FT DOMAIN 1856 1939 FIBRONECTIN TYPE-III 11.
FT DOMAIN 1962 2039 FIBRONECTIN TYPE-III 12.
FT DOMAIN 2069 2150 FIBRONECTIN TYPE-III 13.
FT DOMAIN 2167 2248 FIBRONECTIN TYPE-III 14.
FT DOMAIN 2266 2347 FIBRONECTIN TYPE-III 15.
FT DOMAIN 2365 2446 FIBRONECTIN TYPE-III 16.
FT DOMAIN 2471 2552 FIBRONECTIN TYPE-III 17.
FT DOMAIN 2582 2663 FIBRONECTIN TYPE-III 18.
FT DOMAIN 2688 2769 FIBRONECTIN TYPE-III 19.
FT DOMAIN 2794 2875 FIBRONECTIN TYPE-III 20.
FT DOMAIN 2889 2972 FIBRONECTIN TYPE-III 21.
FT DOMAIN 2997 3078 FIBRONECTIN TYPE-III 22.
FT DOMAIN 3105 3186 FIBRONECTIN TYPE-III 23.
FT DOMAIN 3211 3292 FIBRONECTIN TYPE-III 24.
FT DOMAIN 3307 3384 FIBRONECTIN TYPE-III 25.
FT DOMAIN 3399 3481 FIBRONECTIN TYPE-III 26.
FT DOMAIN 3494 3575 FIBRONECTIN TYPE-III 27.
FT DOMAIN 3601 3682 FIBRONECTIN TYPE-III 28.
FT DOMAIN 3699 3787 FIBRONECTIN TYPE-III 29.
FT DOMAIN 3801 3879 FIBRONECTIN TYPE-III 30.
FT DOMAIN 3890 3971 FIBRONECTIN TYPE-III 31.
FT DOMAIN 3978 4059 FIBRONECTIN TYPE-III 32.
FT DOMAIN 4071 4289 FIBRINOGEN C-TERMINAL.
FT SITE 1748 1750 CELL ATTACHMENT SITE (POTENTIAL).
FT DISULFID 187 197 BY SIMILARITY.
FT DISULFID 191 202 BY SIMILARITY.
FT DISULFID 204 213 BY SIMILARITY.
FT DISULFID 218 228 BY SIMILARITY.
FT DISULFID 222 233 BY SIMILARITY.
FT DISULFID 235 244 BY SIMILARITY.
FT DISULFID 249 259 BY SIMILARITY.
FT DISULFID 253 264 BY SIMILARITY.
FT DISULFID 266 275 BY SIMILARITY.
FT DISULFID 280 290 BY SIMILARITY.
FT DISULFID 284 295 BY SIMILARITY.
FT DISULFID 297 306 BY SIMILARITY.
FT DISULFID 311 321 BY SIMILARITY.
FT DISULFID 315 326 BY SIMILARITY.
FT DISULFID 328 337 BY SIMILARITY.
FT DISULFID 342 352 BY SIMILARITY.
FT DISULFID 346 357 BY SIMILARITY.
FT DISULFID 359 368 BY SIMILARITY.
FT DISULFID 373 383 BY SIMILARITY.
FT DISULFID 377 388 BY SIMILARITY.
FT DISULFID 390 399 BY SIMILARITY.
FT DISULFID 404 414 BY SIMILARITY.
FT DISULFID 408 419 BY SIMILARITY.
FT DISULFID 421 430 BY SIMILARITY.
FT DISULFID 435 445 BY SIMILARITY.
FT DISULFID 439 450 BY SIMILARITY.

Query Match 18.0%; Score 59; DB 1; Length 4289;
Best Local Similarity 31.7%; Pred. No. 1.1e+02;
Matches 20; Conservative 4; Mismatches 15; Indels 24; Gaps 4;

QY 2 QEPVKGVPSTKPGSC-PIILIR-----CAMLNPPNRLKDTDCPGIKKCC 46
Db 681 EEP---PASACPGGCGPRELCRAGQCVCEGFRGPDCAIQTCPG-----DCRGRGECHD 731

QY 47 GSC 49

Db 732 GSC 734

RESULT 52

TFF2_MOUSE TFF2_MOUSE STANDARD; PRT; 129 AA.

AC Q03404; O08913;
DT 01-OCT-1994 (Rel. 30, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Trefoil factor 2 precursor (Spasmolytic polypeptide) (SP).
GN TFF2 OR SML1 OR SP.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE OF 2-129 FROM N.A.
RC TISSUE=Stomach, and Pancreas;
RX MEDLINE=90151615; PubMed=2303034;
RA Tomasetto C., Rio M.C., Gautier C., Wolf C., Hareuveni M., Chambon P.,
RA Lathe R.;
RT "hsp, the domain-duplicated homolog of p52 protein, is co-expressed
with p52 in stomach but not in breast carcinoma.";
RL EMBO J. 9:407-414(1990).
RN [2]

RP SEQUENCE FROM N.A.

RC STRAIN=129/Sv;
RA Taupin D., Bei H., Hocker M., Wang T., Podolsky D.;
RL Submitted (JUN-1997) to the EMBL/GenBank/DBJ databases.
CC -!- FUNCTION: INHIBITS GASTROINTESTINAL MOTILITY AND GASTRIC ACID
SECRETION. COULD FUNCTION AS A STRUCTURAL COMPONENT OF GASTRIC
MUCUS. POSSIBLY BY STABILIZING GLYCOPROTEINS IN THE MUCUS GEL
THROUGH INTERACTIONS WITH CARBOHYDRATE SIDE CHAINS (BY
SIMILARITY).
CC -!- SUBCELLULAR LOCATION: Secreted.
CC -!- TISSUE SPECIFICITY: STOMACH AND PANCREAS.
CC -!- SIMILARITY: CONTAINS 2 P-TYPE (TREFOIL) DOMAINS.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
or send an email to license@isb-sib.ch).

CC EMBL; X51697; CAA35994.1; -.

CC EMBL; U78770; AAB58490.1; -.

CC HSSP; P01359; 1PCP.

CC MGD; MGI:1306805; Tff2.

CC InterPro; IPR000519; P_trefoil.

CC Pfam; PF00088; trefoil; 2.

CC PRINTS; PR00680; PTREFOIL.

CC SMART; SM00018; P; 2.

CC PROSITE; PS00025; P_TREFOIL; 2.

KW Repeat; Signal.

FT SIGNAL 1 23 POTENTIAL.

FT CHAIN 24 129 TREFOIL FACTOR 2.

FT DOMAIN 30 72 P-TYPE 1.

FT DOMAIN 80 121 P-TYPE 2.

FT DISULFID 29 127 BY SIMILARITY.

FT DISULFID 31 58 BY SIMILARITY.
FT DISULFID 42 57 BY SIMILARITY.
FT DISULFID 52 69 BY SIMILARITY.
FT DISULFID 81 107 BY SIMILARITY.
FT DISULFID 91 106 BY SIMILARITY.
FT DISULFID 101 118 BY SIMILARITY.
SQ SEQUENCE 129 AA; 14172 MW; 25A37794199175D4 CRC64;

Query Match 17.7%; Score 58; DB 1; Length 129;

Best Local Similarity 37.8%; Pred. No. 7.8;

Matches 17; Conservative 5; Mismatches 11; Indels 12; Gaps 4;

QY 12 KPGSCPIILIRCAMLNPPNRLKDTDCPGI--KKCCGSGCGMACF 54

Db 25 KPSPC-----RCSRLTPHNR--KNCGFPGITSEQCQFDLGC---CF 59

RESULT 53

ACSA_CRYPV ACSA_CRYPV STANDARD; PRT; 694 AA.

AC Q27549;

DT 30-MAY-2000 (Rel. 39, Created)

DT 30-MAY-2000 (Rel. 39, Last sequence update)

DT 16-OCT-2001 (Rel. 40, Last annotation update)

DE Acetyl-coenzyme A synthetase (EC 6.2.1.1) (Acetate--CoA ligase) (Acyl-
activating enzyme).
GN ACS.

OS Cryptosporidium parvum.

OC Eukaryota; Alveolata; Apicomplexa; Coccidia; Eimeriida;

OC Cryptosporidiidae; Cryptosporidium.

OX NCBI_TaxID=5807;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=KSU-1;

RX MEDLINE=96237870; PubMed=8636846;

RA Khramtsov N.V., Blunt D.S., Montelone B.A., Upton S.J.;

RT "The putative acetyl-CoA synthetase gene of Cryptosporidium parvum
and a new conserved protein motif in acetyl-CoA synthetases.";

RL J. Parasitol. 82:423-427(1996).

CC -!- CATALYTIC ACTIVITY: ATP + acetate + CoA = AMP + diphosphate +
acetyl-CoA.

CC -!- SIMILARITY: BELONGS TO THE ATP-DEPENDENT AMP-BINDING ENZYME
FAMILY.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
or send an email to license@isb-sib.ch).

CC EMBL; U24082; AAC47128.1; -.

CC InterPro; IPR000873; AMP-bind.

CC Pfam; PF00501; AMP-binding; 1.

CC PRINTS; PR00154; AMPBINDING.

CC PROSITE; PS00455; AMP_BINDING; 1.

KW Ligase.

SQ SEQUENCE 694 AA; 77970 MW; 5BA42E16E9AEE63C CRC64;

Query Match 17.7%; Score 58; DB 1; Length 694;

Best Local Similarity 38.6%; Pred. No. 31;

Matches 17; Conservative 4; Mismatches 19; Indels 4; Gaps 2;

QY 4 PVKGPVSTKPGSC--PIILIRCAMLNPPNRLKDTDCPGIKKCC 45

Db 461 PIPGCFDTKPGSATFFFGIEPAILDPTG--KEIDGPGSGVLC 502

RESULT 54

SORL_HUMAN

ID -SORL_HUMAN STANDARD; PRT; 2214 AA.

AC Q92673; Q92856;

FT DISULFID 810 BY SIMILARITY.
FT DISULFID 867 BY SIMILARITY.
FT DISULFID 889 BY SIMILARITY.
FT DISULFID 898 BY SIMILARITY.
FT DISULFID 914 BY SIMILARITY.
FT DISULFID 1060 BY SIMILARITY.
FT DISULFID 1071 BY SIMILARITY.
FT DISULFID 1089 BY SIMILARITY.
FT DISULFID 1153 BY SIMILARITY.
FT DISULFID 1149 BY SIMILARITY.
FT DISULFID 1126 BY SIMILARITY.
FT DISULFID 1196 BY SIMILARITY.
FT DISULFID 1234 BY SIMILARITY.
FT DISULFID 1272 BY SIMILARITY.
FT DISULFID 1669 BY SIMILARITY.
FT DISULFID 1686 BY SIMILARITY.
FT DISULFID 1879 BY SIMILARITY.
FT DISULFID 1899 BY SIMILARITY.
FT DISULFID 1972 BY SIMILARITY.
FT DISULFID 1950 BY SIMILARITY.
FT DISULFID 1927 BY SIMILARITY.
FT DISULFID 1993 BY SIMILARITY.
FT DISULFID 2724 BY SIMILARITY.
FT DISULFID 2739 BY SIMILARITY.
FT DISULFID 2750 BY SIMILARITY.
FT DISULFID 2754 BY SIMILARITY.
FT DISULFID ? BY SIMILARITY.
FT CARBOHYD 99 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 156 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 211 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 666 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 857 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 1231 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 1515 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 1574 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 2223 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 2290 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 2357 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 2400 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 2546 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 2585 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 2790 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CONFLICT 55 G -> E (IN REF. 2).
FT CONFLICT 70 V -> I (IN REF. 3).
FT CONFLICT 266 A -> G (IN REF. 2).
FT CONFLICT 280 I -> V (IN REF. 2).
FT CONFLICT 409 VCH -> ICQ (IN REF. 2).
FT CONFLICT 994 G -> A (IN REF. 1).
FT CONFLICT 1021 F -> L (IN REF. 2).
FT CONFLICT 2381 L -> P (IN REF. 2).
FT CONFLICT 2406 P -> L (IN REF. 2).
SQ SEQUENCE 2813 AA; 309716 MW; 5DF93E1E5E72F80C CRC64;

Query Match 17.7%; Score 58; DB 1; Length 2813;
Best Local Similarity 40.0%; Pred. No. 99;
Matches 18; Conservative 2; Mismatches 9; Indels 16; Gaps 5;

QY 23 CAMLNPP----NRCLKDTCPIKKCEG---SCG----MACFVP 56
Db 2199 CAMSCPPSLVYNHC--EHGCP---RLCEGNTSSCGDQPSEGCFCP 2238

RESULT 56
MT_CRAVI STANDARD; PRT; 74 AA.
AC P23038;
DT 01-NOV-1991 (Rel. 20, Created)
DT 01-APR-1993 (Rel. 25, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Metallothionein (MT).
OS Crassostrea virginica (Eastern oyster).
OC Eukaryota; Metazoa; Mollusca; Bivalvia; Pteriomorpha; Ostreoida;
OC Ostreoida; Ostreidae; Crassostrea.

OX NCBI_TaxID=6565;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=91363394; PubMed=1888750;
RA Unger M.E.; Chen T.T.; Murphy C.M.; Vestling M.M.; Fenselau C.;
RA Roesijadi G.;
RT "Primary structure of molluscan metallothioneins deduced from PCR-amplified cDNA and mass spectrometry of purified proteins.";
RL Biochim. Biophys. Acta 1074:371-377(1991).
RN [2]
RP SEQUENCE OF 1-27.
RC TISSUE=Gill;
RX MEDLINE=89372898; PubMed=2774559;
RA Roesijadi G.; Kielland S.; Klerks P.;
RT "Purification and properties of novel molluscan metallothioneins.";
RL Arch. Biochem. Biophys. 273:403-413(1989).
RN [3]
RP SEQUENCE OF 1-24, AND ACETYLATION.
RC TISSUE=Gill;
RX MEDLINE=91291855; PubMed=2065077;
RA Roesijadi G.; Vestling M.M.; Murphy C.M.; Klerks P.L.; Fenselau C.C.;
RT "Structure and time-dependent behavior of acetylated and non-acetylated forms of a molluscan metallothionein.";
RL Biochim. Biophys. Acta 1074:230-236(1991).
CC -!- FUNCTION: THE METALLOTHIONEINS ARE INVOLVED IN THE CELLULAR SEQUESTRATION OF TOXIC METAL IONS.
CC -!- INDUCTION: BY CADMIUM.
CC -!- PTM: THIS MT OCCURS IN BOTH ACETYLATED AND NON-ACETYLATED FORMS.
CC -!- SIMILARITY: BELONGS TO THE METALLOTHIONEIN SUPERFAMILY; FAMILY 2.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).
CC -----
CC EMBL; X59862; CAA42522.1; -.
DR PIR; S05645; S05645.
DR PIR; S17156; S17156.
DR HSSP; P55949; 1DME.
DR InterPro; IPR003019; Metallthion.
DR InterPro; IPR001008; Mtlthion_misc.
DR Pfam; PF00131; metalthio; 1.
DR PRINTS; PR00875; MTMOLLUSC.
KW Metal-binding; Metal-thiolate cluster; Cadmium; Acetylation.
FT INIT MET 0
FT MOD_RES 1 1 ACETYLATION (PARTIAL).
SQ SEQUENCE 74 AA; 7213 MW; 6364DFB4046E205E CRC64;

Query Match 17.6%; Score 57.5; DB 1; Length 74;
Best Local Similarity 35.4%; Pred. No. 5.6;
Matches 17; Conservative 2; Mismatches 20; Indels 9; Gaps 3;

QY 15 SCPIILIRCAMLNPPNRLKDTDCPIKKCC---EG--SCGMACFVP 56
Db 17 SCPATGCKC---GPGCKGDDCKCAGCKVKCSCTSEGGCKGCKCTGP 61

RESULT 57
ADEN_ADEF1 STANDARD; PRT; 204 AA.
AC Q9IIH4;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Adenain (EC 3.4.22.39) (Endoprotease) (Late L3 23 kDa protein).
OS Frog adenovirus 1.
OC Viruses; dsDNA viruses, no RNA stage; Adenoviridae; Siadenovirus.
OX NCBI_TaxID=114102;
RN [1]
RP SEQUENCE FROM N.A.

RC STRAIN=ATCC VR-896;
RA Davison A.J., Wright K.M., Harrach B.;
RT "Phylogenetic position of an amphibian adenovirus";
RL Submitted (JAN-2000) to the EMBL/GenBank/DBJ databases.
CC -!- FUNCTION: THIOL PROTEASE CLEAVING SPECIFIC GLY-ALA PEPTIDES IN A
CC NUMBER OF VIRAL PRECURSOR PROTEINS (TERMINAL, IIA, VI, VII, VIII,
CC IIA AND MU) WHICH ARE REQUIRED FOR VIRION MATURATION. ALSO CLEAVES
CC HOST CELLS CYTOSKELETAL KERATINS K7 AND K18.
CC -!- CATALYTIC ACTIVITY: Cleaves proteins of the adenovirus and its
CC host cell at two consensus sites: -Yaa-Xaa-Gly-Gly-|-Xaa- and
CC -Yaa-Xaa-Gly-Xaa-|-Gly- (in which Yaa is Met, Ile or Leu, and Xaa
CC is any amino acid).
CC -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY C5.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; AF224336; AAF86933.1; -
CC HSSP; P03252; IAVP.
CC MEROPS; C05.001; -
CC InterPro; IPR000855; Peptidase C5.
CC Pfam; PF00770; Peptidase C5; 1.
CC PRINTS; PR00703; ADVENTOPTASE.
CC ProDom; PD003705; Peptidase C5; 1.
KW Hydrolase; Thiol protease; Late protein.
FT ACT_SITE 54 54 BY SIMILARITY.
FT ACT_SITE 71 71 BY SIMILARITY.
FT ACT_SITE 121 121 BY SIMILARITY.
SQ SEQUENCE 204 AA; 23222 MW; 91DC0B01544577A5 CRC64;

Query Match 17.6%; Score 57.5; DB 1; Length 204;
Best Local Similarity 28.6%; Pred. No. 13;
Matches 10; Conservative 8; Mismatches 16; Indels 1; Gaps 1;

QY 20 LIRCAMLNPPNRLKDTDCPGIKKC-CEGSCGMAC 53
Db 91 ILKSTALTKPSRCIKLVKSKEAVQCTCSAACGLFC 125

RESULT 58
LRP2 RAT
ID LRP2 RAT STANDARD; PRT; 4660 AA.
AC P98158;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Low-density lipoprotein receptor-related protein 2 precursor (Megalin)
DE (Glycoprotein 330).
GN LRP2.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Sprague-Dawley; TISSUE=Kidney;
RX MEDLINE=95024033; PubMed=7937880;
RA Saito A., Pietromonaco S., Loo A.K.C., Farquhar M.G.;
RT "Complete cloning and sequencing of rat gp330/megalin," a
RT distinctive member of the low density lipoprotein receptor gene
RT family.";
RL Proc. Natl. Acad. Sci. U.S.A. 91:9725-9729(1994).
RN [2]
RP FUNCTION.
RX MEDLINE=95386696; PubMed=7544804;
RA Moestrup S.K., Cui S., Vorum H., Bregengaard C., Bjorn S.E.,
RA Norris K., Gliemann J., Christensen E.I.;
RT "Evidence that epithelial glycoprotein 330/megalin mediates uptake of

polybasic drugs.";
J. Clin. Invest. 96:1404-1413(1995).
[3]
TISSUE SPECIFICITY.
RX MEDLINE=94172242; PubMed=7510321;
RA Zheng G., Bachinsky D.R., Stamenkovic I., Strickland D.K., Brown D.,
RA Andres G., McCluskey R.T.;
RT "Organ distribution in rats of two members of the low-density
RT lipoprotein receptor gene family, gp330 and LRP/alpa 2MR, and the
RT receptor-associated protein (RAP).";
RL J. Histochem. Cytochem. 42:531-542(1994).
CC -!- FUNCTION: BINDS PLASMINOGEN, EXTRACELLULAR MATRIX COMPONENTS,
CC PLASMINOGEN ACTIVATOR-PLASMINOGEN, EXTRACELLULAR BETA-VLDL, LIPOPROTEIN LIPASE,
CC COMPLEX, APOLIPOPROTEIN E-ENRICHED BETA-VLDL, LIPOPROTEIN LIPASE,
CC LACTOFERRIN, CLUSTERIN AND CALCIUM.
CC -!- FUNCTION: RECEPTOR-MEDIATED UPTAKE OF POLYBASIC DRUGS SUCH AS
CC APROTININ, AMINOGLYCOSIDES AND POLYMYXIN B.
CC -!- SUBUNIT: FORMS A MULTIMERIC COMPLEX TOGETHER WITH A RECEPTOR-
CC ASSOCIATED PROTEIN (RAP).
CC -!- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN. EXPRESSED IN
CC CLATHRIN-COATED PITS; A SOLUBLE FORM IS POSSIBLY DERIVED BY
CC CLEAVAGE AT THE CELL SURFACE.
CC -!- TISSUE SPECIFICITY: EPITHELIAL CELLS OF KIDNEY GLOMERULUS AND
CC PROXIMAL TUBULE, LUNG, EPIDIDYMIS, YOLK SAC, AMONG OTHERS.
CC -!- SIMILARITY: CONTAINS 36 LDL-RECEPTOR CLASS A DOMAINS.
CC -!- SIMILARITY: CONTAINS 37 LDL-RECEPTOR CLASS B DOMAINS.
CC -!- SIMILARITY: CONTAINS 17 EGF-LIKE DOMAINS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; L34049; AAA51369.1; -
CC HSSP; Q07954; LCR8.
CC GlycoSuiteDB; P98158; -
CC InterPro; IPR000152; Asx hydroxyl.
CC InterPro; IPR000561; EGF-like.
CC InterPro; IPR001881; EGF Ca.
CC InterPro; IPR002172; LDL_recept_A.
CC InterPro; IPR000033; Ldl_receptor_rep.
CC Pfam; PF00057; ldl_recept_a; 36.
CC Pfam; PF00058; ldl_recept_b; 33.
CC PRINTS; PR00261; LDLRECEPTOR.
CC SMART; SM00179; EGF CA; 3.
CC SMART; SM00001; EGF_like; 15.
CC SMART; SM00192; LDLA; 36.
CC SMART; SM00135; LY; 35.
CC PROSITE; PS00010; ASX_HYDROXYL; 4.
CC PROSITE; PS00022; EGF_1; 1.
CC PROSITE; PS01186; EGF_2; 8.
CC PROSITE; PS01187; EGF CA; 3.
CC PROSITE; PS01209; LDLRA_1; 31.
CC PROSITE; PS00068; LDLRA_2; 36.
KW Glycoprotein; Repeat; Endocytosis; Coated pits; Transmembrane;
KW Receptor; EGF-like domain; SH3-binding; Signal.
FT SIGNAL 1 25
FT CHAIN 26 4660
FT LOW-DENSITY LIPOPROTEIN RECEPTOR-RELATED
FT PROTEIN 2.
FT EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 4426 4446
FT DOMAIN 4447 4660
FT DOMAIN 26 64
FT DOMAIN 65 105
FT DOMAIN 106 144
FT DOMAIN 145 181
FT DOMAIN 181 219
FT DOMAIN 220 258
FT DOMAIN 263 307
FT DOMAIN 308 346
FT POTENTIAL.
FT LDL-RECEPTOR CLASS A 1.
FT LDL-RECEPTOR CLASS A 2.
FT LDL-RECEPTOR CLASS A 3.
FT LDL-RECEPTOR CLASS A 4.
FT LDL-RECEPTOR CLASS A 5.
FT LDL-RECEPTOR CLASS A 6.
FT LDL-RECEPTOR CLASS A 7.
FT EGF-LIKE 1.

FT DOMAIN 347 385 EGF-LIKE 2. 385
FT DOMAIN 435 LDL-RECEPTOR CLASS B 1. 435
FT DOMAIN 478 LDL-RECEPTOR CLASS B 2. 478
FT DOMAIN 519 LDL-RECEPTOR CLASS B 3. 519
FT DOMAIN 566 LDL-RECEPTOR CLASS B 4. 566
FT DOMAIN 611 LDL-RECEPTOR CLASS B 5. 611
FT DOMAIN 658 LDL-RECEPTOR CLASS B 6. 658
FT DOMAIN 704 EGF-LIKE 3. 704
FT DOMAIN 752 LDL-RECEPTOR CLASS B 7. 752
FT DOMAIN 795 LDL-RECEPTOR CLASS B 8. 795
FT DOMAIN 837 LDL-RECEPTOR CLASS B 9. 837
FT DOMAIN 881 LDL-RECEPTOR CLASS B 10. 881
FT DOMAIN 969 EGF-LIKE 4. 969
FT DOMAIN 1023 LDL-RECEPTOR CLASS B 11. 1023
FT DOMAIN 1064 LDL-RECEPTOR CLASS B 12. 1064
FT DOMAIN 1108 LDL-RECEPTOR CLASS B 13. 1108
FT DOMAIN 1148 LDL-RECEPTOR CLASS B 14. 1148
FT DOMAIN 1186 LDL-RECEPTOR CLASS B 15. 1186
FT DOMAIN 1225 LDL-RECEPTOR CLASS B 16. 1225
FT DOMAIN 1269 LDL-RECEPTOR CLASS B 17. 1269
FT DOMAIN 1308 LDL-RECEPTOR CLASS B 18. 1308
FT DOMAIN 1351 LDL-RECEPTOR CLASS B 19. 1351
FT DOMAIN 1390 EGF-LIKE 5. 1390
FT DOMAIN 1430 EGF-LIKE 6. CALCIUM-BINDING (POTENTIAL). 1430
FT DOMAIN 1479 LDL-RECEPTOR CLASS B 20. 1479
FT DOMAIN 1522 LDL-RECEPTOR CLASS B 21. 1522
FT DOMAIN 1567 LDL-RECEPTOR CLASS B 22. 1567
FT DOMAIN 1611 LDL-RECEPTOR CLASS B 23. 1611
FT DOMAIN 1656 LDL-RECEPTOR CLASS B 24. 1656
FT DOMAIN 1701 EGF-LIKE 7. 1701
FT DOMAIN 1791 LDL-RECEPTOR CLASS B 25. 1791
FT DOMAIN 1834 LDL-RECEPTOR CLASS B 26. 1834
FT DOMAIN 1884 LDL-RECEPTOR CLASS B 27. 1884
FT DOMAIN 1932 LDL-RECEPTOR CLASS B 28. 1932
FT DOMAIN 1973 LDL-RECEPTOR CLASS B 29. 1973
FT DOMAIN 2019 EGF-LIKE 8. 2019
FT DOMAIN 2108 LDL-RECEPTOR CLASS B 30. 2108
FT DOMAIN 2156 LDL-RECEPTOR CLASS B 31. 2156
FT DOMAIN 2201 LDL-RECEPTOR CLASS B 32. 2201
FT DOMAIN 2245 LDL-RECEPTOR CLASS B 33. 2245
FT DOMAIN 2289 LDL-RECEPTOR CLASS B 34. 2289
FT DOMAIN 2291 EGF-LIKE 9. 2291
FT DOMAIN 2332 LDL-RECEPTOR CLASS B 35. 2332
FT DOMAIN 2384 LDL-RECEPTOR CLASS B 36. 2384
FT DOMAIN 2477 LDL-RECEPTOR CLASS B 37. 2477
FT DOMAIN 2518 LDL-RECEPTOR CLASS B 38. 2518
FT DOMAIN 2562 LDL-RECEPTOR CLASS B 39. 2562
FT DOMAIN 2604 LDL-RECEPTOR CLASS B 40. 2604
FT DOMAIN 2647 LDL-RECEPTOR CLASS B 41. 2647
FT DOMAIN 2694 EGF-LIKE 10. 2694
FT DOMAIN 2738 LDL-RECEPTOR CLASS B 42. 2738
FT DOMAIN 2779 LDL-RECEPTOR CLASS B 43. 2779
FT DOMAIN 2820 LDL-RECEPTOR CLASS B 44. 2820
FT DOMAIN 2862 LDL-RECEPTOR CLASS B 45. 2862
FT DOMAIN 2903 LDL-RECEPTOR CLASS B 46. 2903
FT DOMAIN 2947 LDL-RECEPTOR CLASS B 47. 2947
FT DOMAIN 2992 LDL-RECEPTOR CLASS B 48. 2992
FT DOMAIN 3031 LDL-RECEPTOR CLASS B 49. 3031
FT DOMAIN 3072 LDL-RECEPTOR CLASS B 50. 3072
FT DOMAIN 3112 EGF-LIKE 11. 3112
FT DOMAIN 3153 EGF-LIKE 12. CALCIUM-BINDING (POTENTIAL). 3153
FT DOMAIN 3194 EGF-LIKE 13. 3194
FT DOMAIN 3241 LDL-RECEPTOR CLASS B 51. 3241
FT DOMAIN 3282 LDL-RECEPTOR CLASS B 52. 3282
FT DOMAIN 3333 LDL-RECEPTOR CLASS B 53. 3333
FT DOMAIN 3377 LDL-RECEPTOR CLASS B 54. 3377
FT DOMAIN 3420 LDL-RECEPTOR CLASS B 55. 3420
FT DOMAIN 3461 EGF-LIKE 14. 3461
FT DOMAIN 3511 LDL-RECEPTOR CLASS B 56. 3511
FT DOMAIN 3552 LDL-RECEPTOR CLASS B 57. 3552
FT DOMAIN 3593 LDL-RECEPTOR CLASS B 58. 3593
FT DOMAIN 3634 LDL-RECEPTOR CLASS B 59. 3634
FT DOMAIN 3675 LDL-RECEPTOR CLASS B 60. 3675
FT DOMAIN 3718 LDL-RECEPTOR CLASS B 61. 3718
FT DOMAIN 3759 LDL-RECEPTOR CLASS B 62. 3759
FT DOMAIN 3797 LDL-RECEPTOR CLASS B 63. 3797
FT DOMAIN 3836 LDL-RECEPTOR CLASS B 64. 3836
FT DOMAIN 3882 LDL-RECEPTOR CLASS B 65. 3882

FT DOMAIN 3883 3924 LDL-RECEPTOR CLASS A 35.
FT DOMAIN 3928 3966 LDL-RECEPTOR CLASS A 36.
FT DOMAIN 3968 4008 EGF-LIKE 14.
FT DOMAIN 4009 4050 EGF-LIKE 15. CALCIUM-BINDING (POTENTIAL).
FT DOMAIN 4156 4197 LDL-RECEPTOR CLASS B 35.
FT DOMAIN 4199 4241 LDL-RECEPTOR CLASS B 36.
FT DOMAIN 4244 4284 LDL-RECEPTOR CLASS B 37.
FT DOMAIN 4332 4370 EGF-LIKE 16.
FT DOMAIN 4379 4413 EGF-LIKE 17.
FT SITE 4454 4460 SH3-BINDING (POTENTIAL).
FT SITE 4457 4463 SH3-BINDING (POTENTIAL).
FT SITE 4606 4609 SH2-BINDING (POTENTIAL).
FT SITE 4619 4625 SH3-BINDING (POTENTIAL).
FT SITE 4624 4630 SH3-BINDING (POTENTIAL).
FT SITE 1743 1745 CELL ATTACHMENT SITE (POTENTIAL).
FT SITE 4522 4527 ENDOCYTOSIS SIGNAL (POTENTIAL).
FT SITE 4601 4606 ENDOCYTOSIS SIGNAL (POTENTIAL).
FT DISULFID 28 40 BY SIMILARITY.
FT DISULFID 35 53 BY SIMILARITY.
FT DISULFID 47 62 BY SIMILARITY.
FT DISULFID 67 80 BY SIMILARITY.
FT DISULFID 74 93 BY SIMILARITY.
FT DISULFID 87 103 BY SIMILARITY.
FT DISULFID 108 120 BY SIMILARITY.
FT DISULFID 115 133 BY SIMILARITY.
FT DISULFID 127 142 BY SIMILARITY.
FT DISULFID 147 157 BY SIMILARITY.

Query Match 17.6%; Score 57.5; DB 1; Length 4660;
Best Local Similarity 34.1%; Pred. No. 1.7e+02;
Matches 14; Conservative 4; Mismatches 18; Indels 5; Gaps 2;

QY 14 GSCPILIRCAMLPNRLKDT-DCPGIKKCCGSGCMAC 53
Db 1148 GTCQPTQFRC----PDHRCISPLYVCDGDKDCADGSDGAC 1184

RESULT 59

M3KA_HUMAN
ID M3KA_HUMAN STANDARD; PRT; 954 AA.
AC Q02779; Q12761; Q14871;
DT 01-FEB-1994 (Rel. 28, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Mitogen-activated protein kinase kinase 10 (EC 2.7.1.37)
DE (Mixed lineage kinase 2) (Protein kinase MST).
GN MAP3K10 OR MLK2 OR MST.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Brain;
RX MEDLINE=96128179; PubMed=8536694;
RA Dorow D.S., Devereux L., Tu G.F., Price G., Nicholl J.K.,
RA Sutherland G.R., Simpson R.J.;
RT "Complete nucleotide sequence, expression, and chromosomal
RT localisation of human mixed-lineage kinase 2.";
RL Eur. J. Biochem. 234:492-500(1995).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Brain;
RX MEDLINE=95249256; PubMed=7731697;
RA Kato M., Hirai M., Sugimura T., Terada M.;
RT "Cloning and characterization of MST, a novel (putative)
RL serine/threonine kinase with SH3 domain.";
RN [3]
RP SEQUENCE OF 244-480 FROM N.A.
RC TISSUE=Colon epithelium;
RX MEDLINE=93238756; PubMed=8477742;
RA Dorow D.S., Devereux L., Dietzsch E., de Kretser T.;

RT "Identification of a new family of human epithelial protein kinases
RT containing two leucine/isoleucine-zipper domains.";
RL Eur. J. Biochem. 213:701-710(1993).
CC -!- CATALYTIC ACTIVITY: ATP + a protein = ADP + a phosphoprotein.
CC -!- TISSUE SPECIFICITY: EXPRESSED IN BRAIN AND SKELETAL MUSCLE.
CC -!- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.
CC MAP KINASE KINASE KINASE SUBFAMILY.
CC -!- SIMILARITY: CONTAINS 1 SH3 DOMAIN.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; X90846; CAA62351.1; -;
DR EMBL; Z48615; CAA88531.1; -;
DR PIR; S32468; S32468.
DR HSSP; P11362; 1FGK.
DR Genew; HGNC:6849; MAP3K10.
DR MIM; 600137; -;
DR InterPro; IPR000719; Euk_pkinase.
DR InterPro; IPR001452; SH3.
DR InterPro; IPR004040; STY_pkinase.
DR InterPro; IPR002290; Ser_thr_pkinase.
DR InterPro; IPR001245; Tyr_pkinase.
DR Pfam; PF00018; SH3; 1.
DR Pfam; PF00069; pkinase; 1.
DR PRINTS; PR00452; SH3DOMAIN.
DR PRINTS; PR00109; TYRKINASE.
DR ProDom; PD000001; Euk_pkinase; 1.
DR ProDom; PD000066; SH3; 1.
DR SMART; SM00326; SH3; 1.
DR SMART; SM00221; STYKC; 1.
DR PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
DR PROSITE; PS00108; PROTEIN_KINASE_ST; 1.
DR PROSITE; PS00111; PROTEIN_KINASE_DOM; 1.
DR PROSITE; PS00002; SH3; 1.
DR Transferase; Serine/threonine-protein kinase; Tyrosine-protein kinase;
KW ATP-binding; SH3 domain.
KW DOMAIN 2 5 POLY-GLU.
FT DOMAIN 16 81 SH3.
FT DOMAIN 98 360 PROTEIN KINASE.
FT NP BIND 104 112 ATP (BY SIMILARITY).
FT BINDING 125 125 ATP (BY SIMILARITY).
FT ACT SITE 222 222 BY SIMILARITY.
FT DOMAIN 384 405 LEUCINE-ZIPPER 1 (BY SIMILARITY).
FT DOMAIN 419 440 LEUCINE-ZIPPER 2 (BY SIMILARITY).
FT DOMAIN 449 463 ARG/LYS-RICH (BASIC).
FT CONFLICT 462 464 SRL -> AV (IN REF. 2).
FT CONFLICT 465 480 LKLRGGSHISLPSGF -> AQAGRRQPHQPALWL (IN
FT REF. 3).
FT CONFLICT 471 471 G -> S (IN REF. 2).
FT CONFLICT 807 807 G -> R (IN REF. 2).
FT CONFLICT 818 818 V -> A (IN REF. 2).
SQ SEQUENCE 954 AA; 103623 MW; 538F4AAA559B0ABA CRC64;

Query Match 17.4%; Score 57; DB 1; Length 954;
Best Local Similarity 34.9%; Pred. No. 52;
Matches 15; Conservative 3; Mismatches 13; Indels 12; Gaps 1;

QY 3 EPVKGVPSTKPGSCP-----IILRCAMLNPPNRC 33
DB 130 DPEKDPAVTAEQVCQEARLFGALQHPNFIALRGACLNPPHLC 172

RESULT 60
ID IEL1_MOMCH STANDARD; PRT; 28 AA.
AC P10296;
DT 01-MAR-1989 (Rel. 10, Created)

DT 01-MAR-1989 (Rel. 10, Last sequence update)
DT 01-MAR-1989 (Rel. 10, Last annotation update)
DE Elastase inhibitor I (MCEI-I).
OS Momordica charantia (Bitter melon) (Balsam pear).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eusoids I; Cucurbitales; Cucurbitaceae; Momordica.
OX NCBI_TaxID=3673;
RN [1]
RP SEQUENCE.
RC TISSUE=Seed;
RX MEDLINE=89291812; PubMed=2738047;
RA Hara S., Makino J., Ikenaka T.;
RT "Amino acid sequences and disulfide bridges of serine proteinase
inhibitors from bitter melon (Momordica charantia LINN.) seeds.";
RL J. Biochem. 105:88-92(1989).
CC -!- SIMILARITY: BELONGS TO THE SQUASH FAMILY OF SERINE PROTEASE
CC INHIBITORS.
DR PIR; JX0059; JX0059.
DR HSSP; P01074; 1CTI.
DR InterPro; IPR000737; Squash.
DR Pfam; PF00299; squash; 1.
DR PRINTS; PR00293; SQUASHINHBTR.
DR ProDom; PD003401; Squash; 1.
DR SMART; SM00286; PTI; 1.
DR PROSITE; PS00286; SQUASH_INHIBITOR; 1.
KW Serine protease inhibitor.
FT ACT SITE 5 6 REACTIVE BOND.
FT DISULFID 3 20 BY SIMILARITY.
FT DISULFID 10 22 BY SIMILARITY.
FT DISULFID 16 27 BY SIMILARITY.
SQ SEQUENCE 28 AA; 3169 MW; 961C63D22C74FEC8 CRC64;

Query Match 17.3%; Score 56.5; DB 1; Length 28;
Best Local Similarity 31.4%; Pred. No. 3.2;
Matches 11; Conservative 5; Mismatches 10; Indels 9; Gaps 1;

QY 16 CPILIRCAMLNPPNRCIKDTPGKICKEGSCG 50
DB 3 CPLIWMECK-----RDSCLAQICVDGHC 28

RESULT 61
MT1B_HUMAN STANDARD; PRT; 61 AA.
ID MT1B_HUMAN
AC P07438;
DT 01-APR-1988 (Rel. 07, Created)
DT 01-APR-1988 (Rel. 07, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Metallothionein-IB (MT-IB).
GN MT1B.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Euthera; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=87064506; PubMed=3785191;
RA Heguy A., West A., Richards R.I., Karin M.;
RT "Structure and tissue-specific expression of the human
metallothionein IB gene.";
RL Mol. Cell. Biol. 6:2149-2157(1986).
CC -!- FUNCTION: METALLOTHIONEINS HAVE A HIGH CONTENT OF CYSTEINE
CC RESIDUES THAT BIND VARIOUS HEAVY METALS; THESE PROTEINS ARE
CC TRANSCRIPTIONALLY REGULATED BY BOTH HEAVY METALS AND
CC GLUCOCORTICOID.
CC -!- SUBUNIT: MONOMER.
CC -!- DOMAIN: CLASS I METALLOTHIONEINS CONTAIN 2 METAL-BINDING DOMAINS;
CC FOUR DIVALENT IONS ARE CHELATED WITHIN CLUSTER A OF THE ALPHA
CC DOMAIN AND ARE COORDINATED VIA CYSTEINYL THIOLATE BRIDGES TO 11
CC CYSTEINE LIGANDS. CLUSTER B, THE CORRESPONDING REGION WITHIN THE
CC BETA DOMAIN, CAN LIGATE THREE DIVALENT IONS TO 9 CYSTEINES.
CC -!- SIMILARITY: BELONGS TO THE METALLOTHIONEIN SUPERFAMILY; FAMILY 1.

ID AC FBN1 HUMAN STANDARD; PRT; 2871 AA.
AC P35555;
DT 01-JUN-1994 (Rel. 29, Created)
DT 01-JUN-1994 (Rel. 29, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Fibrillin 1 precursor.
GN FBN1 OR FBN.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OC NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Placenta;
RX MEDLINE=93372860; PubMed=8364578;
RA Pereira L.V., D'Alessio M., Ramirez F., Lynch J.R., Sykes B.,
RA Pangilinan T., Bonadio J.;
RT "Genomic organization of the sequence coding for fibrillin, the
RT defective gene product in Marfan syndrome.";
RL Hum. Mol. Genet. 2:961-968(1993).
RN [2]
RP SEQUENCE OF 1-932 FROM N.A.
RC TISSUE=Placenta, and Fibroblast;
RX MEDLINE=94010947; PubMed=7691719;
RA Corson G.M., Chalberg S.C., Dietz H.C., Charbonneau N.L., Sakai L.Y.;
RT "Fibrillin binds calcium and is coded by cDNAs that reveal a
RT multidomain structure and alternatively spliced exons at the 5'
RT end.";
RL Genomics 17:476-484 (1993).
RN [3]
RP SEQUENCE OF 899-2871 FROM N.A.
RX MEDLINE=91304568; PubMed=1852207;
RA Maslen C.L., Corson G.M., Maddox B.K., Glanville R.W., Sakai L.Y.;
RT "Partial sequence of a candidate gene for the Marfan syndrome.";
RL Nature 352:334-337(1991).
RN [4]
RP SEQUENCE OF 813-1313 FROM N.A.
RX MEDLINE=91304567; PubMed=1852206;
RA Lee B., Godfrey M., Vitale E., Hori H., Mattei M.-G., Sarfarazi M.,
RA Tsipouras P., Ramirez F., Hollister D.W.;
RT "Linkage of Marfan syndrome and a phenotypically related disorder to
RT two different fibrillin genes.";
RL Nature 352:330-334(1991).
RN [5]
RP CHARACTERIZATION.
RX MEDLINE=91317849; PubMed=1860873;
RA Sakai L.Y., Keene D.R., Glanville R.W., Bachinger H.P.;
RT "Purification and partial characterization of fibrillin, a cysteine-
RT rich structural component of connective tissue microfibrils.";
RL J. Biol. Chem. 266:14763-14770(1991).
RN [6]
RP STRUCTURE BY NMR OF 2054-2125.
RX MEDLINE=98031893; PubMed=9362480;
RA Yuan X., Downing A.K., Knott V., Handford P.A.;
RT "Solution structure of the transforming growth factor beta-binding
RT protein-like module, a domain associated with matrix fibrils.";
RL EMBO J. 16:6659-6666(1997).
RN [7]
RP STRUCTURE BY NMR OF 2124-2205.
RX MEDLINE=96144829; PubMed=8568869;
RA Knott V., Downing A.K., Cardy C.M., Handford P.A.;
RT "Calcium binding properties of an epidermal growth factor-like domain
RT pair from human fibrillin-1.";
RL J. Mol. Biol. 255:22-27(1996).
RN [8]
RP STRUCTURE BY NMR OF 2124-2205.
RX MEDLINE=96222301; PubMed=8653794;
RA Downing A.K., Knott V., Werner J.M., Cardy C.M., Campbell I.D.,
RA Handford P.A.;
RT "Solution structure of a pair of calcium-binding epidermal growth
RT factor-like domains: implications for the Marfan syndrome and other
RT genetic disorders.";
RL Cell 85:597-605(1996).

RN [9]
RP REVIEW ON MFS VARIANTS.
RX MEDLINE=96174615; PubMed=8594563;
RA Colod G., Beroud C., Soussi T., Junien C., Boileau C.;
RT "Software and database for the analysis of mutations in the human
RT FBN1 gene.";
RL Nucleic Acids Res. 24:137-141(1996).
RN [10]
RP REVIEW ON MFS VARIANTS.
RX MEDLINE=97169383; PubMed=9016526;
RA Colod-Beroud G., Beroud C., Ades L., Black C., Boxer M., Brock D.J.,
RA Godfrey M., Hayward C., Karttunen L., Milewicz D., Peltonen L.,
RA Richards R.I., Wang W., Junien C., Boileau C.;
RT "Marfan Database (second edition): software and database for the
RT analysis of mutations in the human FBN1 gene.";
RL Nucleic Acids Res. 25:147-150(1997).
RN [11]
RP REVIEW ON VARIANTS.
RX MEDLINE=98062175; PubMed=9401003;
RA Hayward C., Brock D.J.H.;
RT "Fibrillin-1 mutations in Marfan syndrome and other type-1
RT fibrillinopathies.";
RL Hum. Mutat. 10:415-423(1997).
RN [12]
RP VARIANT MFS PRO-1137.
RX MEDLINE=91304569; PubMed=1852208;
RA Dietz H.C., Cutting G.R., Pyeritz R.E., Maslen C.L., Sakai L.Y.,
RA Corson G.M., Puffenberger E.G., Hamosh A., Nanthakumar E.J.,
RA Currstin S.M., Stetten G., Meyers D.A., Francomano C.A.;
RT "Marfan syndrome caused by a recurrent de novo missense mutation in
RT the fibrillin gene.";
RL Nature 352:337-339(1991).
RN [13]
RP VARIANTS MFS SER-1249; ARG-1663; SER-2221 AND SER-2307.
RX MEDLINE=93250834; PubMed=1301946;
RA Dietz H.C., Saraiva J.M., Pyeritz R.E., Cutting G.R., Francomano C.A.;
RT "Clustering of fibrillin (FBN1) missense mutations in Marfan syndrome
RT patients at cysteine residues in EGF-like domains.";
RL Hum. Mutat. 1:366-374(1992).
RN [14]
RP VARIANT MFS SER-2307.
RX MEDLINE=92235290; PubMed=1569206;
RA Dietz H.C., Pyeritz R.E., Puffenberger E.G., Kendzior R.J. Jr.,
RA Corson G.M., Maslen C.L., Sakai L.Y., Francomano C.A., Cutting G.R.;
RT "Marfan phenotype variability in a family segregating a missense
RT mutation in the epidermal growth factor-like motif of the fibrillin
RT gene.";
RL J. Clin. Invest. 89:1674-1680(1992).
RN [15]
RP VARIANTS MFS ILE-548 AND ALA-723.
RX MEDLINE=94010946; PubMed=8406497;
RA Dietz H.C., McIntosh I., Sakai L.Y., Corson G.M., Chalberg S.C.,
RA Pyeritz R.E., Francomano C.A.;
RT "Four novel FBN1 mutations: significance for mutant transcript level
RT and EGF-like domain calcium binding in the pathogenesis of Marfan
RT syndrome.";
RL Genomics 17:468-475(1993).
RN [16]
RP VARIANTS MFS SER-2144.
RX MEDLINE=93278402; PubMed=8504310;
RA Hewett D.R., Lynch J.R., Smith R., Sykes B.C.;
RT "A novel fibrillin mutation in the Marfan syndrome which could
RT disrupt calcium binding of the epidermal growth factor-like module.";
RL Hum. Mol. Genet. 2:475-477(1993).
RN [17]
RP VARIANTS MFS R-862; Y-1117; P-1137 AND F-1589, AND VARIANT A-1148.
RX MEDLINE=94108431; PubMed=8281141;
RA Tynan K., Comeau K., Pearson M., Wilgenbus P., Levitt D., Gasner C.,
RA Berg M.A., Miller D.C., Francke U.;
RT "Mutation screening of complete fibrillin-1 coding sequence: report
RT of five new mutations, including two in 8-cysteine domains.";
RL Hum. Mol. Genet. 2:1813-1821(1993).
RN [18]

RP VARIANTS MFS GLY-217 AMD ARG-2627.
RX MEDLINE=95067970; PubMed=7977366;
RA Karttunen L., Raghunath M., Loennqvist L., Peltonen L.;
RT "A compound-heterozygous Marfan patient: two defective fibrillin
alleles result in a lethal phenotype.";
RL Am. J. Hum. Genet. 55:1083-1091(1994).
RN [19]
RP VARIANT EL LYS-2447.
RX MEDLINE=94245249; PubMed=8188302;
RA Lonnqvist L., Child A., Kainulainen K., Davidson R., Puhakka L.,
RA Peltonen L.;
RT "A novel mutation of the fibrillin gene causing ectopia lentis.";
RL Genomics 19:573-576(1994).
RN [20]
RP VARIANT MFS CYS-627.
RX MEDLINE=94272487; PubMed=8004112;
RA Hayward C., Rae A.L., Porteous M.E.M., Logie L.J., Brock L.J.;
RT "Two novel mutations and a neutral polymorphism in EGF-like domains
of the fibrillin gene (FBN1): SSCP screening of exons 15-21 in Marfan
syndrome patients.";
RL Hum. Mol. Genet. 3:373-375(1994).
RN [21]
RP VARIANT MFS CYS-122.
RX MEDLINE=94314977; PubMed=8040326;
RA Stahl-Hallengren C., Ukkonen T., Kainulainen K., Kristofersson U.,
RA Saxne T., Tornqvist K., Peltonen L.;
RT "An extra cysteine in one of the non-calcium-binding epidermal growth
factor-like motifs of the FBN1 polypeptide is connected to a novel
variant of Marfan syndrome.";
RL J. Clin. Invest. 94:709-713(1994).
RN [22]
RP VARIANT MFS TYR-1223.
RX MEDLINE=94351682; PubMed=8071963;
RA Hewett D.R., Lynch J.R., Child A., Sykes B.C.;
RT "A new missense mutation of fibrillin in a patient with Marfan
syndrome.";
RL J. Med. Genet. 31:338-339(1994).
RN [23]
RP VARIANT MFS HIS-1170.
RX MEDLINE=95174777; PubMed=7870075;
RA Hayward C., Porteous M.E.M., Brock D.J.H.;
RT "A novel mutation in the fibrillin gene (FBN1) in familial
arachnodactyly.";
RL Mol. Cell. Probes 8:325-327(1994).
RN [24]
RP VARIANTS MFS G-217; N-1023; R-1074; Y-1242; R-1513; E-2127; W-2151;
K-2447 AND R-2511.
RX MEDLINE=94184368; PubMed=8136837;
RA Kainulainen K., Karttunen L., Puhakka L., Sakai L., Peltonen L.;
RT "Mutations in the fibrillin gene responsible for dominant ectopia
lentis and neonatal Marfan syndrome.";
RL Nat. Genet. 6:64-69(1994).
RN [25]
RP VARIANT SER-1127.

Query Match 17.1%; Score 56; DB 1; Length 2871;
Best Local Similarity 26.6%; Pred. No. 1.7e+02;
Matches 17; Conservative 8; Mismatches 23; Indels 16; Gaps 4;

QY 8 PVSTKPGSCPII--LIRCAMLNPPNRCCLKDTCPIKKCCE-----GSC-----GMAC 53
Db 1981 PRKCAPGTQCNLDGSRVC--ICPPGYSLQNEKCEDIDECVEEPEICALGTCSNTEGSFKC 2038
QY 54 FVPQ 57
Db 2039 LCPE 2042

RESULT 69
HSTN_VIBCH STANDARD; PRT; 78 AA.
ID HSTN_VIBCH
AC P04429;
DT 13-AUG-1987 (Rel. 05, Created)

DT 01-AUG-1992 (Rel. 23, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Heat-stable enterotoxin ST precursor (Nonagglutinating cholera
vibrios ST) (NAG-ST) (Non O1-ST).
GN STN.
OS Vibrio cholerae, and
OS Vibrio mimicus.
OC Bacteria; Proteobacteria; gamma subdivision; Vibrionaceae; Vibrio.
OX NCBI_TaxID=666, 674;
RN [1]
RP SEQUENCE FROM N.A.
RC SPECIES=V.cholerae; STRAIN=NRT36;
RX MEDLINE=90382953; PubMed=2205577;
RA Ogawa A., Kato J.I., Watanabe H., Nair B.G., Takeda T.;
RT "Cloning and nucleotide sequence of a heat-stable enterotoxin gene
from Vibrio cholerae non-O1 isolated from a patient with traveler's
diarrhea.";
RL Infect. Immun. 58:3325-3329(1990).
RN [2]
RP SEQUENCE OF 62-78.
RC SPECIES=V.cholerae; STRAIN=Serovar non-O1;
RX MEDLINE=86056320; PubMed=4065341;
RA Takao T., Shimonishi Y., Kobayashi M., Nishimura O., Arita M.,
RA Takeda T., Honda T., Miwatani T.;
RT "Amino acid sequence of heat-stable enterotoxin produced by Vibrio
cholerae non-O1";
RL FEBS Lett. 193:250-254(1985).
RN [3]
RP SEQUENCE OF 62-78.
RC SPECIES=V.mimicus; STRAIN=AQ-0915;
RX MEDLINE=91257548; PubMed=2044934;
RA Arita M., Honda T., Miwatani T., Takeda T., Takao T., Shimonishi Y.;
RT "Purification and characterization of a heat-stable enterotoxin of
Vibrio mimicus";
RL FEMS Microbiol. Lett. 63:105-110(1991).
CC -!- SIMILARITY: BELONGS TO THE HEAT-STABLE ENTEROTOXIN FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
or send an email to license@isb-sib.ch).
CC -----
CC EMBL; M85198; AAA64889.1; -
DR PIR; A41469; QHVC1.
DR HSSP; P01559; 1ETN.
DR InterPro; IPR001489; Enterotoxin_HS.
DR Pfam; PF02048; Enterotoxin_HS; 1.
DR PROSITE; PS00273; ENTEROTOXIN_H_STABLE; 1.
KW Toxin; Enterotoxin; Signal.
FT SIGNAL 1 18 POTENTIAL.
FT PROPEP 19 61
FT CHAIN 62 78 HEAT-STABLE ENTEROTOXIN ST.
FT DISULFID 64 69 BY SIMILARITY.
FT DISULFID 65 73 BY SIMILARITY.
FT DISULFID 68 76 BY SIMILARITY.
SQ SEQUENCE 78 AA; 8830 MW; EED385B298D0964F CRC64;

Query Match 17.0%; Score 55.5; DB 1; Length 78;
Best Local Similarity 36.7%; Pred. No. 9.7;
Matches 11; Conservative 4; Mismatches 6; Indels 9; Gaps 1;

QY 34 KDTDCPGIKK-----CCGSCGMACF 54
Db 45 ENEECPFIKQVDENGNLIDCCEICCNPA CF 74

RESULT 70
HSTO_VIBCH STANDARD; PRT; 78 AA.
ID HSTO_VIBCH
AC Q07425;


```
CC SIMILARITY).
CC -!- ALTERNATIVE PRODUCTS: 4 ISOFORMS; 1/RBCK1 (SHOWN HERE), 2/RBCK2, 3
CC AND 4; MAY BE PRODUCED BY ALTERNATIVE SPLICING.
CC -!- PTM: PHOSPHORYLATED (BY SIMILARITY).
CC -!- SIMILARITY: CONTAINS 1 UBIQUITIN-LIKE DOMAIN.
CC -!- SIMILARITY: CONTAINS 1 RING-TYPE ZINC FINGER.
CC -!- SIMILARITY: CONTAINS 1 B BOX-TYPE ZINC FINGER.
CC -!- SIMILARITY: CONTAINS 1 RANBP2-TYPE ZINC FINGER.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; U67322; AAD00162.1; -.
CC EMBL; AL121747; CAC17516.1; -.
CC EMBL; AL121747; CAC28311.1; -.
CC EMBL; AL121747; CAC28312.1; -.
CC EMBL; BC000983; AAH00983.3; -.
CC Genew; HGNC:15864; C20orf18.
CC InterPro; IPR000626; Ubiquitin.
CC InterPro; IPR000315; Znf_RanGDP.
CC InterPro; IPR001876; Znf_RanGDP.
CC InterPro; IPR001841; Znf_Ring.
CC Pfam; PF00097; zf-C3HC4; 1.
CC Pfam; PF00641; zf-RanBP; 1.
CC SMART; SM00184; RING; 1.
CC SMART; SM00547; Znf_RBZ; 1.
CC PROSITE; PS50053; UBIQUITIN_2; 1.
CC PROSITE; PS50119; ZF_BBOX; FALSE NEG.
CC PROSITE; PS01358; ZF_RANBP2_1; 1.
CC PROSITE; PS50199; ZF_RANBP2_2; 1.
CC PROSITE; PS00518; ZF_RING_1; 1.
CC PROSITE; PS50089; ZF_RING_2; 1.
CC Ub1 conjugation pathway; Zinc-finger; Coiled coil; Phosphorylation;
KW alternative splicing.
FT DOMAIN 45 109 UBIQUITIN-LIKE.
FT ZN_FING 183 212 RANBP2-TYPE.
FT DOMAIN 223 251 COILED COIL (POTENTIAL).
FT ZN_FING 272 317 RING-TYPE.
FT ZN_FING 366 401 B BOX-TYPE.
FT ZN_FING 437 463 B BOX-LIKE-TYPE.
FT VARSPLIC 1 45 MALSLTRAVAGGDEQVAMKCAIWLAEQVRVLSVLKPEVSP
FT TQDI -> MGTATPDGREDQE (IN ISOFORM 3 AND
FT ISOFORM 4).
FT VARSPLIC 243 262 RKQQQEGNYLQHVLDQRS -> GVPAGHHPPQPGGGGLL
FT PLH (IN ISOFORM 2 AND ISOFORM 4).
FT VARSPLIC 263 500 MISSING (IN ISOFORM 2 AND ISOFORM 4).
FT SEQUENCE 500 AA; 56381 MW; 7492E03EB2E9CE4F CRC64;

Query Match 17.0%; Score 55.5; DB 1; Length 500;
Best Local Similarity 30.0%; Pred. No. 45;
Matches 15; Conservative 3; Mismatches 15; Indels 17; Gaps 3;

QY 2 QEPVKGVPSTKP-----GSCPIILIRCAMLNPPNRCCLKDTCDCPGIKKCC 45
ID VG47_HSVI1 STANDARD; PRT; 518 AA.
AC Q00139;
DT 01-DEC-1992 (Rel. 24, Created)
DT 01-DEC-1992 (Rel. 24, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE 56 kDa serine protease (EC 3.4.21.-).
GN 47.
OS Ictalurid herpesvirus 1 (Channel catfish virus) (CCV).
OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
```

```
OC Ictalurid Herpes-like viruses.
OX NCBI_TaxID=10401;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Auburn 1;
RX MEDLINE=92087490; PubMed=1727613;
RA Davison A.J.;
RT "Channel catfish virus: a new type of herpesvirus.";
RL Virology 186:9-14(1992).
CC -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S8.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; M75136; AAA88150.1; -.
CC PIR; C36791; PRBE11.
CC MEROPS; S08.0PA; -.
CC InterPro; IPR002884; P_domain.
CC InterPro; IPR000209; Peptidase_S8.
CC Pfam; PF01483; P; PARTIAL.
CC ProDom; PD000717; P_domain; 1.
CC PROSITE; PS00136; SUBTILASE_ASP; FALSE NEG.
CC PROSITE; PS00137; SUBTILASE_HIS; FALSE NEG.
CC PROSITE; PS00138; SUBTILASE_SER; FALSE NEG.
KW Hydrolase; Serine protease.
FT ACT_SITE 334 334
FT SEQUENCE 518 AA; 55984 MW; DAAAF919A975460D CRC64;

Query Match 17.0%; Score 55.5; DB 1; Length 518;
Best Local Similarity 37.5%; Pred. No. 46;
Matches 15; Conservative 5; Mismatches 13; Indels 7; Gaps 3;

QY 18 IILIRCAMLNPPNRCCLKDTCDCPGIKKCC--CE---GSCG 50
ID ITR4_CUCMA STANDARD; PRT; 32 AA.
AC P07853;
DT 01-AUG-1988 (Rel. 08, Created)
DT 01-MAR-1989 (Rel. 10, Last sequence update)
DT 01-MAY-1992 (Rel. 22, Last annotation update)
DE Trypsin inhibitors IV and III (CMTI-IV and CMTI-III).
OS Cucurbita maxima (Pumpkin) (Winter squash).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eurosids I; Cucurbitales; Cucurbitaceae; Cucurbita.
OX NCBI_TaxID=3661;
RN [1]
RP SEQUENCE (CMT-IV).
RC TISSUE=Seed;
RX MEDLINE=85149300; PubMed=3977882;
RA Wiczorek M., Otlewski J., Cook J., Parks K., Leluk J.,
RA Wilimowska-Pelc A., Polanowski A., Wilusz T., Laskowski M. Jr.;
RT "The squash family of serine proteinase inhibitors. Amino acid
RT sequences and association equilibrium constants of inhibitors from
RT squash, summer squash, zucchini, and cucumber seeds.";
RL Biochem. Biophys. Res. Commun. 126:646-652(1985).
RN [2]
RP SEQUENCE (CMT-III).
RC TISSUE=Seed;
RX MEDLINE=83184077; PubMed=6840699;
RA Wilusz T., Wiczorek M., Polanowski A., Denton A., Cook J.,
RA Laskowski M. Jr.;
RT "Amino-acid sequence of two trypsin isoinhibitors, ITD I and ITD III
RT from squash seeds (Cucurbita maxima).";
```

RL Hoppe-Seyler's Z. Physiol. Chem. 364:93-95(1983).
RN [3]
RP PRELIMINARY SEQUENCE (CMT-III).
RC TISSUE=Seed;
RX MEDLINE=82005824; PubMed=7275008;
RA Nowak K., Slominska A., Polanowski A., Wieczorek M., Wilusz T.;
RT "Trypsin inhibitor III from squash seeds (Cucurbita maxima), its
RT reactive site and proposed amino acid sequence.";
RL Hoppe-Seyler's Z. Physiol. Chem. 362:1017-1019(1981).
RN [4]
RP STRUCTURE BY NMR OF CMT-III.
RX MEDLINE=92118901; PubMed=1731946;
RA Krishnamoorthi R., Gong Y., Lin C.-L.S., Vandervelde D.;
RT "Two-dimensional NMR studies of squash family inhibitors. Sequence-
RT specific proton assignments and secondary structure of reactive-site
RT hydrolyzed Cucurbita maxima trypsin inhibitor III.";
RL Biochemistry 31:898-904(1992).
CC -!- SIMILARITY: BELONGS TO THE SQUASH FAMILY OF SERINE PROTEASE
CC INHIBITORS.
DR PIR; A01313; TIPU3.
DR PIR; S07156; S07156.
DR HSSP; P01074; ICTI.
DR InterPro; IPR000737; Squash.
DR Pfam; PF00299; squash; 1.
DR PRINTS; PRO0293; SQUASHINHBTR.
DR ProDom; PD003401; Squash; 1.
DR SMART; SM00286; FTI; 1.
DR PROSITE; PS00286; SQUASH_INHIBITOR; 1.
KW Serine protease inhibitor.
FT CHAIN 1 32 CMTI-IV.
FT CHAIN 4 32 CMTI-III.
FT ACT SITE 8 9 REACTIVE BOND.
FT DISULFID 6 23 BY SIMILARITY.
FT DISULFID 13 25 BY SIMILARITY.
FT DISULFID 19 31 BY SIMILARITY.
SQ SEQUENCE 32 AA; 3669 MW; 0F591120B0137512 CRC64;

Query Match 16.8%; Score 55; DB 1; Length 32;
Best Local Similarity 38.9%; Pred. No. 5.3;
Matches 14; Conservative 3; Mismatches 9; Indels 10; Gaps 2;

Qy 16 CPIILIRCAMLNPPNRCLKDTDCPGIKKCCCE-GSCG 50
|||:::| ||:| | | | |
Db 6 CPRILMKCK-----KSDCLAEVCVLEHGYCG 32

Search completed: February 11, 2003, 08:37:03
Job time : 16 secs

GenCore version 5.1.4_p5_4578
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: February 15, 2003, 19:03:02 ; Search time 1295 Seconds
(without alignments)
1280.973 Million cell updates/sec

Title: 09-833799-13B
Perfect score: 327
Sequence: 1 aqepvkgpvtkpgscpiil.....cpgikkccgscgmactvpq 57

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 2054640 seqs, 14551402878 residues

Total number of hits satisfying chosen parameters: 4109280

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 100 summaries

Command line parameters:
-MODEL=frame+p2n.model -DEV=xlp
-Q=/cgn2_1/USPTO spool/US09833799/runat_12022003_120042_23635/app_query.fasta_1.199
-DB=GenEmbl -QFMT=fastap -SUFFIX=rge -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=1 -MATRIX=biosum62 -TRANS=human40.cdi -LIST=100
-DOCALIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=75 -MODE=LOCAL
-OUTFMT=pto -NORM=ext -HEAPSIZ=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US09833799@cgn_1_1_3745 @runat_12022003_120042_23635 -NCPU=6 -ICPU=3
-NO_XLPXY -NO_MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -LONGLOG -DEV_TIMEOUT=120
-WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : GenEmbl:
1: gb_ba:*
2: gb_htg:*
3: gb_in:*
4: gb_om:*
5: gb_ov:*
6: gb_pat:*
7: gb_ph:*
8: gb_pl:*
9: gb_pr:*
10: gb_ro:*
11: gb_sts:*
12: gb_sy:*
13: gb_un:*
14: gb_vi:*
15: em_ba:*
16: em_fun:*
17: em_hum:*
18: em_in:*
19: em_mu:*
20: em_om:*
21: em_or:*
22: em_ov:*
23: em_pat:*
24: em_ph:*
25: em_pl:*
26: em_ro:*
27: em_sts:*
28: em_un:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	327	100.0	171	6	A31088 DNA fragmen
2	327	100.0	206	6	A31086 DNA fragmen
C 3	327	100.0	206	6	A31087 DNA fragmen
4	327	100.0	347	6	A31089 DNA fragmen
C 5	327	100.0	434	6	AX041087 Sequence
6	327	100.0	478	9	HSANTLEUP
C 7	327	100.0	480	6	AX397178 Sequence
8	327	100.0	504	6	A31090 DNA fragmen
9	327	100.0	571	6	I74316 Sequence 14
10	327	100.0	582	9	BC010952 Homo sapi
11	327	100.0	1878	9	HUMELAFIN
12	327	100.0	2309	6	AX333644 Sequence
13	327	100.0	2309	6	AX334316 Sequence
14	327	100.0	2309	9	HUMPREELAS
15	327	100.0	2309	9	S58717
16	327	100.0	99747	9	HSJ172H20
17	324	99.1	177	6	I96047 Sequence 3
18	324	99.1	8598	6	E16036 cDNA encodi
19	323	98.8	177	6	I96048 Sequence 5
20	323	98.8	177	6	I96049 Sequence 7
21	322	98.5	321	6	I74313 Sequence 11
22	250	76.5	573	4	BOT223216 Bos tauru
23	221.5	67.7	3693	4	PIGWAPA
24	217.5	66.5	1128	4	AB003281 Phacochoe
25	216.5	66.2	464	4	PIGWAPD
26	216.5	66.2	1034	4	D83668 Sus scrofa
27	201	61.5	1169	4	AB003283 Pecari ta
28	182	55.7	270	9	MAM223215 Macaca mu
29	181.5	55.5	1325	4	AB003282 Phacochoe
30	179.5	54.9	483	4	PIGSPA12S2
31	179.5	54.9	737	6	E06782 cDNA contai
32	179.5	54.9	789	4	DMY245
33	179.5	54.9	3782	4	PIGWAPB
34	177.5	54.3	578	4	PIGWAPE
35	171	52.3	490	10	AF178426 Rattus no
36	170	52.0	495	4	AB011010 Bos tauru
37	166	50.8	734	4	AB003284 Hippopota
38	165	50.5	600	4	PIGALP
39	161	49.2	671	10	MMU88093 Mus musculu
40	161	49.2	682	10	MMU94341 Mus muscu
41	161	49.2	894	10	BC028509 Mus muscu
42	161	49.2	1123	10	MMU73004 Rattus no
43	160.5	49.1	667	10	AF151982 Rattus no
44	160	48.9	3670	4	PIGWAPC
45	158	48.3	464	10	AB042257 Cavia por
46	156	47.7	194143	2	AC112730 Rattus no
47	148	45.3	2435	10	AF002719 Mus muscu
C 48	148	45.3	190669	10	AL590429 Mouse DNA
49	144	44.0	4090	10	AF205374 Mus muscu
50	143	43.7	194	6	AR067991 Sequence

```
51 143 43.7 324 6 I08404
52 143 43.7 399 6 AX302535
53 143 43.7 399 9 AF114471 Homo sapi
54 143 43.7 498 6 I01501
55 143 43.7 565 6 AX069252
56 143 43.7 565 6 I01502
57 143 43.7 573 9 HSSLIPR
58 143 43.7 594 6 AX328399
59 143 43.7 594 6 AX334507
60 143 43.7 594 6 AX335376
61 143 43.7 594 9 HSLALPR
62 143 43.7 599 6 AX014898
63 143 43.7 625 9 BC020708
64 143 43.7 63796 9 HS30012
65 141 43.1 2657 9 HSSLIPG
66 140 42.8 321 6 A08111
67 138 42.2 321 6 A08112
68 137 41.9 321 6 A08107
69 137 41.9 321 6 A08108
70 137 41.9 321 6 A08110
71 135 41.3 74 6 A31082
72 135 41.3 321 6 A08109
73 135 41.3 1054 4 AB003285
74 131 40.1 321 6 A08106
75 131 40.1 5109 10 AB058645
76 128 39.1 321 6 A08105
77 127 38.8 60989 2 AC094865
78 125 38.2 68 6 A31080
79 124 37.9 71 6 A31083
80 123 37.6 69 6 A31081
81 121.5 37.2 963 9 AY038181
82 118.5 36.2 126270 9 HS211D12
83 117 35.8 36189 3 AF036687
84 116.5 35.6 11008 3 AF078161
85 116.5 35.6 190549 10 AL591512
86 115 35.2 1805 5 AB076019
87 113.5 34.7 1043 10 BC019734
88 113 34.6 19753 3 CBG21N17
89 112 34.3 762 6 AX403457
90 112 34.3 774 9 AY037803
91 110.5 33.8 1370 5 SFU67854
92 110.5 33.8 211873 10 AL596122
93 110.5 33.8 216184 2 AC055773
94 108.5 33.2 471 5 CHKXAL
95 108 33.0 291 4 BOT223217
96 108 33.0 330 4 BOT223218
97 107 32.7 411 6 AX334505
98 107 32.7 411 6 AX335289
99 107 32.7 411 6 AX335639
100 106 32.4 6110 9 HSM801421
```

ALIGNMENTS

```
RESULT 1
A31088
LOCUS A31088 171 bp DNA linear PAT 21-AUG-1995
DEFINITION DNA fragment from patent EP0402068.
ACCESSION A31088
VERSION A31088.1 GI:1249301
KEYWORDS synthetic construct.
SOURCE synthetic construct.
ORGANISM artificial sequences.
REFERENCE 1 (bases 1 to 171)
AUTHORS Christophers,E., Schroder,J.M., Pioli,D., Wiedow,O. and Edge,M.D.
TITLE Polypeptides and polypeptide analogues with inhibitory activity
JOURNAL Patent: EP 0402068-A 29 12-DEC-1990;
IMPERIAL CHEMICAL INDUSTRIES PLC
FEATURES Location/Qualifiers
source 1..171
/organism="synthetic construct"
```

```
BASE COUNT 36 a 51 c 42 g 40 t 2 others
ORIGIN
Alignment Scores: 7.26e-29 Length: 171
Pred. No.: 327.00 Matches: 57
Score: 100.00% Conservative: 0
Percent Similarity: 100.00% Mismatches: 0
Best Local Similarity: 100.00% Indels: 0
Query Match: 100.00% Gaps: 0
DB: 6
09-833799-13B (1-57) x A31088 (1-171)
Qy 1 AlaGlnGluProValLysGlyProValSerThrLysProGlySerCysProIleLeu 20
Db 1 GCGCAAGAGCCAGTCAAGGTCAGTCTCCACTAAGCCTGGCTCCTGCCCCATATCTTG 60
Qy 21 IleArgCysAlaMetLeuAsnProProAsnArgCysLeuLysAspThrAspCysProGly 40
Db 61 ATCCGTGCGCCATGTTGAATCCCCCTAACCCGCTGCTTGAAAGATACTGACTGCCAGGA 120
Qy 41 IleLysLysCysCysGluGlySerCysGlyMetAlaCysPheValProGln 57
Db 121 ATHAAGAARTGCTGTGAAGGCTCTTGGGGATGGCCTGTTTCGTTCCCCAG 171
```

```
RESULT 2
A31086
LOCUS A31086 206 bp DNA linear PAT 21-AUG-1995
DEFINITION DNA fragment from patent EP0402068.
ACCESSION A31086
VERSION A31086.1 GI:1249300
KEYWORDS synthetic construct.
SOURCE synthetic construct.
ORGANISM artificial sequences.
REFERENCE 1 (bases 1 to 206)
AUTHORS Christophers,E., Schroder,J.M., Pioli,D., Wiedow,O. and Edge,M.D.
TITLE Polypeptides and polypeptide analogues with inhibitory activity
JOURNAL Patent: EP 0402068-A 27 12-DEC-1990;
IMPERIAL CHEMICAL INDUSTRIES PLC
FEATURES Location/Qualifiers
source 1..206
/organism="synthetic construct"
```

```
BASE COUNT 51 a 46 c 44 g 65 t
ORIGIN
Alignment Scores: 8.82e-29 Length: 206
Pred. No.: 327.00 Matches: 57
Score: 100.00% Conservative: 0
Percent Similarity: 100.00% Mismatches: 0
Best Local Similarity: 100.00% Indels: 0
Query Match: 100.00% Gaps: 0
DB: 6
09-833799-13B (1-57) x A31086 (1-206)
Qy 1 AlaGlnGluProValLysGlyProValSerThrLysProGlySerCysProIleLeu 20
Db 30 GCTCAAGAACCCAGTTAAAGTCTGTGTCTACTAAGCCAGGTTCTTGTCTATCTTG 89
Qy 21 IleArgCysAlaMetLeuAsnProProAsnArgCysLeuLysAspThrAspCysProGly 40
Db 90 ATTCGTTGCGCTATGTTAAACCCACCTAACCGTTGTTTGAAGGACACTGATGTCAGGT 149
Qy 41 IleLysLysCysCysGluGlySerCysGlyMetAlaCysPheValProGln 57
Db 150 ATCAAAAAGTGTGTGAAGGTTCTCGGGATGGCTTGTTCGTTCCACAA 200
```

```
RESULT 3
A31087/c
```


REFERENCE 1 (bases 1 to 478)
AUTHORS Molhuizen,H.O., Alkemade,H.A., Zeeuwen,P.L., de Jongh,G.J.,
Wieringa,B. and Schalkwijk,J.
TITLE SKAP/ela1: an elastase inhibitor from cultured human
keratinocytes. Purification, cDNA sequence, and evidence for
transglutaminase cross-linking
J. Biol. Chem. 268 (16), 12028-12032 (1993)
JOURNAL 93280175
MEDLINE 7685029
PUBMED
REFERENCE 2 (bases 1 to 478)
AUTHORS Molhuizen,H.O.F.
TITLE Direct Submission
JOURNAL Submitted (20-NOV-1992) Molhuizen H.O.F., Academic Hospital,
Dermatology, Nijmegen, The Netherlands, 6500 HB
FEATURES
source
1. .478
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="pGESKA"
/cell_type="keratinocyte"
/clone_lib="HL1110b Clontech"
1. .478
/partial
/citation={1}
3. .356
/function="proteinase inhibitor"
/citation={1}
/codon_start=1
/evidence=experimental
/product="skin-derived antileukoproteinase"
/protein_id="CAA79223.1"
/db_xref="GI:28712"
/db_xref="SWISS-PROT:P19957"
/translation="MRASSFLIVVFLIAGTLVLEAAVTGVPVKQDITVKGVPFNGQ
DPVKGVSVKQDKVKAQEPVKGPVSTKPGSCPILIRCAMLNPPNRLKDTDCPGIK
KCEGSGCMACFVPQ"
sig_peptide 3. .68
mat_peptide 69. .353
/partial
/product="skin-derived antileukoproteinase"
/function="proteinase inhibitor"
BASE COUNT 102 a 142 c 119 g 115 t
ORIGIN
Alignment Scores:
Pred. No.: 2.13e-28 Length: 478
Score: 327.00 Matches: 57
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 9 Gaps: 0
09-833799-13B (1-57) x HSANTLEUP (1-478)
QY 1 AlaGlnGluProValLysGlyProValSerThrLysProGlySerCysProIleLeu 20
Db 183 GCGCAAGAGCCAGTCAAAGGTCAGTCTCCACTAAGCCTGGCTCCTGCCCATTTATCTTG 242
QY 21 IleArgCysAlaMetLeuAsnProProAsnArgCysLeuLysAspThrAspCysProGly 40
Db 243 ATCCGGTGCCTGCTGTAAGTCCCTTAACCGCTGTTGAAAGATACTGACTGCCCAGGA 302
QY 41 IleLysLysCysCysGluGlySerCysGlyMetAlaCysPheValProGln 57
Db 303 ATCAAGAAGTGCTGTGAAGGCTCTTGCGGGATGGCTGTTTCGTTCCCCAG 353
RESULT 7
AX397178/c AX397178 480 bp DNA linear PAT 18-MAY-2002
LOCUS AX397178
DEFINITION Sequence 1393 from Patent WO0212328.
ACCESSION AX397178
VERSION AX397178.1 GI:21067925
KEYWORDS

SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1
AUTHORS King,G.E., Meagher,M.J., Xu,J. and Secrist,H.
TITLE Compositions and methods for the therapy and diagnosis of colon
cancer
JOURNAL Patent: WO 0212328-A 1393 14-FEB-2002;
CORIXA CORPORATION (US)
FEATURES
source
1. .480
/organism="Homo sapiens"
/db_xref="taxon:9606"
BASE COUNT 113 a 105 c 139 g 120 t 3 others
ORIGIN
Alignment Scores:
Pred. No.: 2.14e-28 Length: 480
Score: 327.00 Matches: 57
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 6 Gaps: 0
09-833799-13B (1-57) x AX397178 (1-480)
QY 1 AlaGlnGluProValLysGlyProValSerThrLysProGlySerCysProIleLeu 20
Db 369 GCGCAAGAGCCAGTCAAAGGTCAGTNTCCACTAAGCCTGGCTCCTGCCCATTTATCTTG 310
QY 21 IleArgCysAlaMetLeuAsnProProAsnArgCysLeuLysAspThrAspCysProGly 40
Db 309 ATCCGGTGCCTGCTGTAAGTCCCTTAACCGCTGTTGAAAGATACTGACTGCCCAGGA 250
QY 41 IleLysLysCysCysGluGlySerCysGlyMetAlaCysPheValProGln 57
Db 249 ATCAAGAAGTGCTGTAAGGCTCTTGCGGGATGGCTGTTTCGTTCCCCAG 199
RESULT 8
AX31090
LOCUS AX31090 504 bp DNA linear PAT 21-AUG-1995
DEFINITION DNA fragment from patent EP0402068.
ACCESSION AX31090
VERSION AX31090.1 GI:1249303
KEYWORDS
SOURCE synthetic construct.
ORGANISM synthetic construct
artificial sequences.
REFERENCE 1 (bases 1 to 504)
AUTHORS Christophers,E., Schroder,J.M., Pioli,D., Wiedow,O. and Edge,M.D.
TITLE Polypeptides and polypeptide analogues with inhibitory activity
against human elastase
JOURNAL Patent: EP 0402068-A 31 12-DEC-1990;
IMPERIAL CHEMICAL INDUSTRIES PLC
FEATURES
source
1. .504
/organism="synthetic construct"
/db_xref="taxon:32630"
BASE COUNT 110 a 152 c 120 g 122 t
ORIGIN
Alignment Scores:
Pred. No.: 2.25e-28 Length: 504
Score: 327.00 Matches: 57
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 6 Gaps: 0
09-833799-13B (1-57) x AX31090 (1-504)
QY 1 AlaGlnGluProValLysGlyProValSerThrLysProGlySerCysProIleLeu 20

|||||
Db 158 GCGCAAGAGCCAGTCAAAGGTCAGTCTCCACTAAGGCTGGCTCCTGCCCATATCTTG 217
QY 21 IleArgCysAlaMetLeuAsnProProAsnArgCysLeuLysAspThrAspCysProGly 40
Db 218 ATCCGGTGCAGCCATGTTGAATCCCTTAACCGCTGCTGAAAGATACTGACTGCCAGGA 277
QY 41 IleLysLysCysCysGluGlySerCysGlyMetAlaCysPheValProGln 57
Db 278 ATCAAGAAGTGTGTGAAGGCTCTTGGGGATGGCTGTTTCGTTCCCCAG 328
RESULT 9
I74316
LOCUS I74316 174316 571 bp DNA linear PAT 03-APR-1998
DEFINITION Sequence 14 from patent US 5688641.
ACCESSION I74316
VERSION I74316.1 GI:3010457
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 571)
AUTHORS Sager,R., Zou,Z., Lee,S.Whan. and Tomasetto,C.Laure.
TITLE Cancer diagnosis using nucleic acid hybridization
JOURNAL Patent: US 5688641-A 14 18-NOV-1997;
FEATURES Location/Qualifiers
source 1..571
BASE COUNT 138 a 167 c 130 g 136 t
ORIGIN
Alignment Scores:
Pred. No.: 2.57e-28 Length: 571
Score: 327.00 Matches: 57
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: Gaps: 0
09-833799-13B (1-57) x I74316 (1-571)
QY 1 AlaGlnGluProValLysGlyProValSerThrLysProGlySerCysProIleLeu 20
Db 199 GCGCAAGAGCCAGTCAAAGGTCAGTCTCCACTAAGCCTGGCTCCTGCCCATATCTTG 258
QY 21 IleArgCysAlaMetLeuAsnProProAsnArgCysLeuLysAspThrAspCysProGly 40
Db 259 ATCCGGTGCAGCCATGTTGAATCCTCTTAACCGCTGCTTGAAGATACTGACTGCCAGGA 318
QY 41 IleLysLysCysCysGluGlySerCysGlyMetAlaCysPheValProGln 57
Db 319 ATCAAGAAGTGTGTGAAGGCTCTTGGGGATGGCTGTTTCGTTCCCCAG 369
RESULT 10
BC010952
LOCUS BC010952 582 bp mRNA linear PRI 25-JUL-2001
DEFINITION Homo sapiens, Similar to protease inhibitor 3, skin-derived (SKALP), clone MGC:13613 IMAGE:4083155, mRNA, complete cds.
ACCESSION BC010952
VERSION BC010952.1 GI:15012094
KEYWORDS MGC.
SOURCE Homo sapiens.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 582)
AUTHORS Strausberg,R.
TITLE Direct Submission
JOURNAL Submitted (23-JUL-2001) National Institutes of Health, Mammalian Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA
REMARK NIH-MGC Project URL: <http://mgc.nci.nih.gov>

COMMENT

Contact: MGC help desk
Email: cgapbs-remail.nih.gov
Tissue Procurement: ATCC
CDNA Library Preparation: CLONTECH Laboratories, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Institute for Systems Biology
<http://www.systemsbio.org>
contact: amadansystemsbiology.org
Anup Madan, Rachel Dickhoff, Jessica Fahey, Stephanie Ford, Julia Greene, Mark Ketteman and Anuradha Madan

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>
Series: IRAL Plate: 19 Row: d Column: 7
This clone was selected for full length sequencing because it passed the following selection criteria: matched mRNA gi: 4505786.

FEATURES

source

Location/Qualifiers
1..582
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="MGC:13613 IMAGE:4083155"
/tissue_type="Brain, glioblastoma"
/clone_lib="NIH MGC_57"
/lab_host="DH10B"
/note="Vector: pDNR-LIB"
26..379
/codon_start=1
/product="Similar to protease inhibitor 3, skin-derived (SKALP)"
/protein_id="AAH10952.1"
/db_xref="GI:15012095"
/translation="MRASSFLIVVFLIAGTLVLEAAVTGVPVKQDTPVGRVPFNGQDPVKGVSVKQDKVKAQEPVKGPVSTKPGSCPIILIRCAMLNPPNRCLKDTDCPGIKKCCEGSCGMACFVPQ"
BASE COUNT 148 a 166 c 134 g 134 t
ORIGIN

Alignment Scores:

Pred. No.: 2.62e-28 Length: 582
Score: 327.00 Matches: 57
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: Gaps: 0

09-833799-13B (1-57) x BC010952 (1-582)

QY 1 AlaGlnGluProValLysGlyProValSerThrLysProGlySerCysProIleLeu 20

Db 206 GCGCAAGAGCCAGTCAAAGGTCAGTCTCCACTAAGCCTGGCTCCTGCCCATATCTTG 265

QY 21 IleArgCysAlaMetLeuAsnProProAsnArgCysLeuLysAspThrAspCysProGly 40

Db 266 ATCCGGTGCAGCCATGTTGAATCCTCTTAACCGCTGCTTGAAGATACTGACTGCCAGGA 325

QY 41 IleLysLysCysCysGluGlySerCysGlyMetAlaCysPheValProGln 57

Db 326 ATCAAGAAGTGTGTGAAGGCTCTTGGGGATGGCTGTTTCGTTCCCCAG 376

RESULT 11

HUMELAFIN

LOCUS HUMELAFIN 1878 bp DNA linear PRI 14-APR-2000
DEFINITION Human gene for elafin, complete cds.
ACCESSION D13156

VERSION D13156.1 GI:219614

KEYWORDS elafin; elastase inhibitor.

SOURCE Homo sapiens (library: lambda EMEL) DNA, clone lambda hI-G1.

ORGANISM

Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 1878)

AUTHORS

Saheki,T., Ito,F., Hagiwara,H., Saito,Y., Kuroki,J., Tachibana,S. and Hirose,S.

TITLE Primary structure of the human elafin precursor preproelafin deduced from the nucleotide sequence of its gene and the presence of unique repetitive sequences in the prosegment

JOURNAL Biochem. Biophys. Res. Commun. 185 (1), 240-245 (1992)

MEDLINE 92287100

REFERENCE 2 (bases 1 to 1878)

AUTHORS Hirose, S.

TITLE Direct Submission

JOURNAL Submitted (01-SEP-1992) Shigehisa Hirose, Tokyo Institute of Technology, Department of Biological Sciences; Ookayama, Meguro-ku, Tokyo 152, Japan (Tel:03-3726-1111(ex.2226), Fax:03-3729-0335)

COMMENT Submitted (01-SEP-1992) to DDBJ by: Shigehisa Hirose

Dept. of Biological Sciences
Tokyo Institute of Technology
Ookayama, Meguroku
Tokyo 152
Japan
Phone: 03-3726-1111
Fax: 03-3729-0335.

FEATURES
source
1. .1878
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_lib="lambda EMBL"
127. .130
153. .158
181. .325
/number=1
join(247. .325,1185. .1459)
/codon_start=1
/product="preproelafin"
/protein_id="BAA02441.1"
/db_xref="GI:219615"
/translation="MRASSFLIVVFLIAGTLVLEAAVTGVPVKQDQTVKGRVFPNGQDPVKQSVKQDKVKAQEPVKGPVSTKPGSCPILIRCAMLNPNNRCLKDTDCPGIKKCEGSCGMACFVPQ"
250. .324
1286. .1456
/product="elafin"
1185. .1460
/number=2
1695. .1852
/number=3
polyA_signal 1847. .1852
BASE COUNT 473 a 464 c 491 g 450 t
ORIGIN

Alignment Scores:
Pred. No.: 8.96e-28 Length: 1878
Score: 327.00 Matches: 57
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: Gaps: 0

09-833799-13B (1-57) x HUMELAFIN (1-1878)

QY 1 AlaGlnGluProValLysGlyProValSerThrLysProGlySerCysProIleIleLeu 20
|||||
Db 1286 GCGCAAGAGCCAGTCAAGGTCAGTCTCCACTAAGCCTGGCTCCTGCCCATATCTTG 1345
|
QY 21 IleArgCysAlaMetLeuAsnProAsnArgCysLeuLysAspThrAspCysProGly 40
|||||
Db 1346 ATCCGGTGGCCCATGTTGAATCCCCCTAACCGCTGCTTGAAAGATACACTGATGCCCCAGGA 1405
|
QY 41 IleLysLysCysCysGluGlySerCysGlyMetAlaCysPheValProGln 57
|||||
Db 1406 ATCAAGAAGTGCTGTGAAGGCTCTTGGGGATGGCCTGTTTCGTTCCCCAG 1456
|

RESULT 12
AX333644
LOCUS AX333644 2309 bp DNA linear PAT 09-JAN-2002

Alignment Scores:
Pred. No.: 1.11e-27 Length: 2309
Score: 327.00 Matches: 57
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: Gaps: 0

09-833799-13B (1-57) x AX333644 (1-2309)

QY 1 AlaGlnGluProValLysGlyProValSerThrLysProGlySerCysProIleIleLeu 20
|||||
Db 1554 GCGCAAGAGCCAGTCAAGGTCAGTCTCCACTAAGCCTGGCTCCTGCCCATATCTTG 1613
|
QY 21 IleArgCysAlaMetLeuAsnProAsnArgCysLeuLysAspThrAspCysProGly 40
|||||
Db 1614 ATCCGGTGGCCCATGTTGAATCCCCCTAACCGCTGCTTGAAAGATACACTGATGCCCCAGGA 1673
|
QY 41 IleLysLysCysCysGluGlySerCysGlyMetAlaCysPheValProGln 57
|||||
Db 1674 ATCAAGAAGTGCTGTGAAGGCTCTTGGGGATGGCCTGTTTCGTTCCCCAG 1724
|

RESULT 13
AX334316
LOCUS AX334316 2309 bp DNA linear PAT 09-JAN-2002

DEFINITION Sequence 4825 from Patent WO0194629.
ACCESSION AX334316
VERSION AX334316
KEYWORDS AX334316.1 GI:18125035
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarhini; Hominidae; Homo.
REFERENCE 1
AUTHORS Young, P.E., Augustus, M., Carter, K.C., Ebner, R., Endress, G., Horrigan, S., Soppet, D.R. and Weaver, Z.
TITLE Cancer gene determination and therapeutic screening using signature gene sets
JOURNAL Patent: WO 0194629-A 4825 13-DEC-2001;
FEATURES
source
1. .2309
Location/Qualifiers
/organism="Homo sapiens"
/db_xref="taxon:9606"

BASE COUNT 595 a 537 c 603 g 572 t 2 others
ORIGIN

Alignment Scores:
Pred. No.: 1.11e-27 Length: 2309
Score: 327.00 Matches: 57
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: Gaps: 0

DEFINITION Sequence 4153 from Patent WO0194629.
ACCESSION AX333644
VERSION AX333644.1 GI:18124363
KEYWORDS human.
SOURCE Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarhini; Hominidae; Homo.
REFERENCE 1
AUTHORS Young, P.E., Augustus, M., Carter, K.C., Ebner, R., Endress, G., Horrigan, S., Soppet, D.R. and Weaver, Z.
TITLE Cancer gene determination and therapeutic screening using signature gene sets
JOURNAL Patent: WO 0194629-A 4153 13-DEC-2001;
FEATURES
source
1. .2309
Location/Qualifiers
/organism="Homo sapiens"
/db_xref="taxon:9606"

BASE COUNT 595 a 537 c 603 g 572 t 2 others
ORIGIN

Alignment Scores:
Pred. No.: 1.11e-27 Length: 2309
Score: 327.00 Matches: 57
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: Gaps: 0

CAAT_signal
TATA_signal
exon
CDS
sig_peptide
mat_peptide
exon
exon
polyA_signal
BASE COUNT 473 a 464 c 491 g 450 t
ORIGIN

Qy 1 AlaGlnGluProValLysGlyProValSerThrLysProGlySerCysProIleLeu 20
|||||
Db 1554 GCGCAAGAGCCAGTCAAAGGTCCAGTCTCCACTAAGCTGGCTCCTGCCCATATCTTG 1613

Qy 21 IleArgCysAlaMetLeuAsnProProAsnArgCysLeuLysAspThrAspCysProGly 40
|||||
Db 1614 ATCCGGTGGCCCATGTTGAATCCCCCTAACCGCTGCTTGAAGATATACTGACTGCCCAGGA 1673

Qy 41 IleLysLysCysCysGluGlySerCysGlyMetAlaCysPheValProGln 57
|||||
Db 1674 ATCAAGAAGTGCTGTGAAGGCTTTCGGGGATGGCCTGTTTCGTTCCCCAG 1724

RESULT 16
HSJ172H20
LOCUS
DEFINITION
Human DNA sequence from clone RPl-172H20 on chromosome 20q12-13.12
Contains SEMG1 and SEMG2 genes for semenogelin I and II, the PI3
gene for protease inhibitor 3, skin-derived (SKALP), ESTs, STSs and
GSSs, complete sequence.
AL049767
AL049767.12 GI:11345541
KEYWORDS
HTG; PI3; SEMG1; SEMG2; SKALP.
SOURCE
human.
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 99747)
Sehra,H.
Direct Submission
Submitted (29-MAR-2001) Sanger Centre, Hinxton, Cambridgeshire,
CB10 1SA, UK. E-mail enquiries: humquery@sanger.ac.uk Clone
requests: clonerequest@sanger.ac.uk
On Nov 27, 2000 this sequence version replaced gi:10198636.
During sequence assembly data is compared from overlapping clones.
Where differences are found these are annotated as variations
together with a note of the overlapping clone name. Note that the
variation annotation may not be found in the sequence submission
corresponding to the overlapping clone, as we submit sequences with
only a small overlap as described above.
The following abbreviations are used to associate primary accession
numbers given in the feature table with their source databases:
Em:, EMBL; Sw:, SWISSPROT; Tr:, TREMBL; Wp:, WORMPEP; Information
on the WORMPEP database can be found at
http://www.sanger.ac.uk/Projects/C_elegans/wormpep This sequence
was generated from part of bacterial clone contigs of human
chromosome 20, constructed by the Sanger Centre Chromosome 20
Mapping Group. Further information can be found at
http://www.sanger.ac.uk/HGP/Chr20
IMPORTANT: This sequence is not the entire insert of clone
RPl-172H20 It may be shorter because we sequence overlapping
sections only once, except for a 100 base overlap.
The true right end of clone RPl-172H20 is at 99747 in this
sequence. The true left end of clone RPl-30012 is at 28594 in this
sequence. The true right end of clone RPl-211D12 is at 100 in this
sequence. This sequence was finished as follows unless otherwise
noted: all regions were either double-stranded or sequenced with an
alternate chemistry or covered by high quality data (i.e., phred
quality >= 30); an attempt was made to resolve all sequencing
problems, such as compressions and repeats; all regions were
covered by at least one plasmid subclone or more than one M13
subclone; and the assembly was confirmed by restriction digest.
RPl-172H20 is from the library RPCI-1 constructed by the group of
Pieter de Jong. For further details see
http://www.chori.org/bacpac/home.htm
VECTOR: pCYPAC2.
Location/Qualifiers
1..99747
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="20"
/map="q12-13.12"
/clone="RPl-172H20"
/clone_lib="RPCI-1"

misc_feature
complement(1..100)
/note="match: STS: Em:HS211D12S"
1116..1993
/note="L1MB1 repeat: matches 5294..6159 of consensus"
2001..2042
/note="21 copies 2 mer ta 83% conserved"
2272..2376
/note="L1 repeat: matches 5193..5300 of consensus"
2377..2408
/note="16 copies 2 mer ta 84% conserved"
2411..2446
/note="18 copies 2 mer at 80% conserved"
complement(2469..2951)
/note="match: GSS: Em:AQ809877"
2581..2789
/note="MER20 repeat: matches 7..218 of consensus"
2982..3563
/note="match: GSS: Em:AQ386436"
2989..3136
/note="L1M4 repeat: matches 3850..3998 of consensus"
3780..4090
/note="AluSx repeat: matches 1..312 of consensus"
4294..4592
/note="AluSx repeat: matches 1..309 of consensus"
4965..5002
/note="19 copies 2 mer gt 94% conserved"
5291..5487
/note="HAL1 repeat: matches 1501..1692 of consensus"
5569..6048
/note="L1MB1 repeat: matches 5696..6172 of consensus"
6049..6158
/note="MER20 repeat: matches 109..218 of consensus"
6160..6228
/note="L2 repeat: matches 2021..2087 of consensus"
6286..6466
/note="L2 repeat: matches 2329..2497 of consensus"
6548..6731
/note="L1MC4 repeat: matches 7650..7849 of consensus"
6735..8399
/note="L1MA8 repeat: matches 4484..6153 of consensus"
8460..8581
/note="FLAM A repeat: matches 1..122 of consensus"
8589..10634
/note="L1M3e repeat: matches 22..1826 of consensus"
10656..10979
/note="MER74A repeat: matches 248..557 of consensus"
11205..11377
/note="AluSg/x repeat: matches 136..308 of consensus"
11408..11859
/note="LTR16B repeat: matches 2..461 of consensus"
11973..12160
/note="Charlie2 repeat: matches 3534..3727 of consensus"
14238..14548
/note="match: STS: Em:G08051"
14337..14396
/note="30 copies 2 mer ct 75% conserved"
14492..17136
/note="HERVL repeat: matches 4..2665 of consensus"
17140..17589
/note="MLT2B repeat: matches 1..444 of consensus"
17590..17899
/note="L1 repeat: matches 4299..4610 of consensus"
17898..18682
/note="L1M1 repeat: matches -1390..-578 of consensus"
19077..19213
/note="L1ME repeat: matches 5694..5824 of consensus"
20761..20831
/note="L1M4b repeat: matches -259..-184 of consensus"
20853..21014
/note="L1MA9 repeat: matches 6102..6270 of consensus"
21044..21573
/note="L1M4b repeat: matches -263..330 of consensus"
21634..21829

repeat_region /note="L1M4b repeat: matches 330. .524 of consensus" 22891. .23199

repeat_region /note="AluSx repeat: matches 1. .309 of consensus" 25460. .25629

repeat_region /note="LTR16A repeat: matches 201. .378 of consensus" 27612. .27639

repeat_region /note="14 copies 2 mer ac 100% conserved" 27655. .27986

repeat_region /note="L1M3a repeat: matches 202. .531 of consensus" 28959. .29336

repeat_region /note="L1M1 repeat: matches 469. .859 of consensus" 29337. .29711

repeat_region /note="THE1C repeat: matches 1. .371 of consensus" 29712. .31317

repeat_region /note="THE1C-internal repeat: matches 1. .1580 of consensus" 31318. .31729

repeat_region /note="THE1C repeat: matches 1. .371 of consensus" 31730. .31886

repeat_region /note="L1M1 repeat: matches 859. .987 of consensus" 33209. .34261

repeat_region /note="L1M3d repeat: matches 859. .1833 of consensus" 34232. .34942

repeat_region /note="L1M1 repeat: matches 1560. .2657 of consensus" 34941. .36117

repeat_region /note="L1 repeat: matches 3935. .5142 of consensus" 36118. .36412

repeat_region /note="AluJo repeat: matches 9. .292 of consensus" 36430. .37302

repeat_region /note="L1 repeat: matches 2905. .3780 of consensus" 37299. .37425

repeat_region /note="L1 repeat: matches 4912. .5038 of consensus" 37434. .37469

repeat_region /note="L1P repeat: matches 2850. .2885 of consensus" 38377. .39183

repeat_region /note="L1M4c repeat: matches 1312. .1666 of consensus" 39261. .42088

repeat_region /note="L1PA15 repeat: matches 671. .3485 of consensus" 42057. .43098

repeat_region /note="L1PA5 repeat: matches 5104. .6143 of consensus" 43119. .45775

repeat_region /note="L1PA15 repeat: matches 3467. .6157 of consensus" 45840. .45956

repeat_region /note="L1ME repeat: matches 5530. .5647 of consensus" 46081. .46114

repeat_region /note="17 copies 2 mer tc 82% conserved" 46492. .46761

repeat_region /note="L1ME repeat: matches 5521. .5793 of consensus" 46828. .46915

repeat_region /note="L1ME3A repeat: matches 5300. .5388 of consensus" 46916. .47210

repeat_region /note="AluSq repeat: matches 1. .305 of consensus" 47211. .47769

repeat_region /note="L1ME3A repeat: matches 5388. .5937 of consensus" complement(48203. .48496)

misc_feature /note="match: GSS: Em:AQ275745" 48638. .48679

repeat_region /note="21 copies 2 mer ac 97% conserved" 49535. .51104

gene /gene="PI3"

mRNA join(49535. .49615,50475. .50750,50984. .51104) /gene="PI3"

/product="dJ172H20.3 (protease inhibitor 3, skin-derived (SKALP))"

/note="match: cDNAs: Em:Z18538 Em:AJ223218 Em:AJ223216 match: ESTs: Em:AA527557 Em:AA583567 Em:AA582866"

/evidence=not_experimental

join(49537. .49615,50475. .50749)

CDS

Alignment Scores:

Pred. No.: 5.78e-26 Length: 99747

Score: 327.00 Matches: 57

Percent Similarity: 100.00% Conservative: 0

Best Local Similarity: 100.00% Mismatches: 0

Query Match: 100.00% Indels: 0

DB: 9 Gaps: 0

09-833799-13B (1-57) x HSJ172H20 (1-99747)

QY 1 AlaGlnGluProValLysGlyProValSerThrLysProGlySerCysProIleLeu 20

Db 50576 GCGCAAGAGCCAGTCAAAGGTCAGTCTCCACTAAGCCTGGCTCCTGCCCATATCTTG 50635

QY 21 IleArgCysAlaMetLeuAsnProProAsnArgCysLeuLysAspThrAspCysProGly 40

Db 50636 ATCCGGTGGCCATGTTGAATCCCCCTAACCGCTGCTTGAAGATACTGACTGCCCAGGA 50695

QY 41 IleLysLysCysCysGluGlySerCysGlyMetAlaCysPheValProGln 57

Db 50696 ATCAAGAAGTCTGTGAAGGCTCTTGGGGATGGCCTGTTTCGTTCCCCAG 50746

RESULT 17

LOCUS I96047

DEFINITION Sequence 3 from patent US 5734014.

ACCESSION I96047

VERSION I96047.1 GI:3940517

KEYWORDS

SOURCE Unknown.

ORGANISM Unclassified.

REFERENCE 1 (bases 1 to 177)

AUTHORS Ishima,Y., Okawa,N., Yoshida,M., Amagaya,S. and Kaj,i,A.

TITLE Elafin derivative

JOURNAL Patent: US 5734014-A 3 31-MAR-1998;

FEATURES

source Location/Qualifiers

1..177

/organism="unknown"

BASE COUNT 39 a 49 c 48 g 41 t

ORIGIN

Alignment Scores:

Pred. No.: 1.65e-28 Length: 177

Score: 324.00 Matches: 56

Percent Similarity: 100.00% Conservative: 1

Best Local Similarity: 98.25% Mismatches: 0

Query Match: 99.08% Indels: 0

DB: 6 Gaps: 0

09-833799-13B (1-57) x I96047 (1-177)

QY 1 AlaGlnGluProValLysGlyProValSerThrLysProGlySerCysProIleLeu 20

Db 1 GCACAGGAACCAAGTTAAAGTCCGGTGCAGACCAACCCGGCTCTTGCCCGATTATCCTG 60

QY 21 IleArgCysAlaMetLeuAsnProProAsnArgCysLeuLysAspThrAspCysProGly 40

Db 61 ATCCGGTGGCTTTGCTGAACCCCGCAACCGTTGCTGTCTGAAAGACACTGACTGCCCGGT 120

QY 41 IleLysLysCysCysGluGlySerCysGlyMetAlaCysPheValProGln 57

Db 121 ATCAAAAATGCTGCGAAGGTTCTTGGCGGTATGGCATGCTTCTGTTCCGCAG 171

RESULT 18

LOCUS E16036

DEFINITION cDNA encoding improved Pichia elafin.

ACCESSION E16036

VERSION E16036.1 GI:5710719

KEYWORDS JP 199812792-A/1.

SOURCE Pichia pastoris.

ORGANISM Pichia pastoris

REFERENCE Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes; Saccharomycetales; Saccharomycetaceae; Pichia.

AUTHORS 1 (bases 1 to 8598)

Taniyama,M., Yamamoto,T., Okawa,N. and Zushi,M.

TITLE ELAFINS-EXPRESSION VECTOR AND PRODUCTION OF ELAFINS BY UTILIZING THE SAME
JOURNAL Patent: JP 1998127292-A 1 19-MAY-1998;
TSMURA & CO
COMMENT OS Pichia pastoris
PN JP 1998127292-A/1
PD 19-MAY-1998
PF 31-OCT-1996 JP 1996304233
PI TANIYAMA MITSUE, YAMAMOTO TAKASHI, OKAWA NORIYUKI, PI ZUSHI
MAKOTO
PC C12N15/09, C07H21/04, C07K14/39, C12N1/19, C12P21/02, (C12N1/19, PC C12R1:84),
PC (C12P21/02, C12R1:84);
CC strandedness: Double;
CC topology: Linear;
CC hypothetical: No;
CC anti-sense: No; Location/Qualifiers
FH Key
FH source 1..8598
FT /organism='Pichia pastoris'
FT /clone='ppIC9/ELF25L'
FT 949..1377
FT /product='improved Pichia elafin' FT
FT sig_peptide 949..1203
FT mat_peptide 1204..1374
FT /product='improved Pichia elafin' FT
FT misc_feature replace(1276..1278,'atg')
FT /note='native Pichia elafin'.
FEATURES
source Location/Qualifiers
1..8598
/organism="Pichia pastoris"
/db_xref="taxon:4922"
BASE COUNT 2306 a 2025 c 1920 g 2347 t
ORIGIN
Alignment Scores:
Pred. No.: 9.69e-27 Length: 8598
Score: 324.00 Matches: 56
Percent Similarity: 100.00% Conservative: 1
Best Local Similarity: 98.25% Mismatches: 0
Query Match: 99.08% Indels: 0
DB: 6 Gaps: 0
09-833799-13B (1-57) x E16036 (1-8598)
QY 1 AlaGlnGluProValLysGlyProValSerThrLysProGlySerCysProIleLeu 20
Db 1204 GCTCAAGAACCAAGTTAAGGTCGGGTTCGACCAACCGGGCTCTTGGCCGATTATCTG 1263
QY 21 IleArgCysAlaMetLeuAsnProProAsnArgCysLeuLysAspThrAspCysProGly 40
Db 1264 ATCCGCTGCGCTTGTCTGAACCGCGCAACCGTGTGTCTGAAAGACACTGACTGCCCGGGT 1323
QY 41 IleLysLysCysCysGluGlySerCysGlyMetAlaCysPheValProGln 57
Db 1324 ATCAAAAAATGCTGCGAAGGTTCTTGGCGTATGGCATGCTTCGTTCCGCAG 1374
RESULT 19
LOCUS I96048 177 bp DNA linear PAT 01-DEC-1998
DEFINITION Sequence 5 from patent US 5734014.
ACCESSION I96048
VERSION I96048.1 GI:3940518
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 177)
AUTHORS Ishima,Y., Okawa,N., Yoshida,M., Amagaya,S. and Kaj,i,A.
TITLE Elafin derivative
JOURNAL Patent: US 5734014-A 5 31-MAR-1998;
FEATURES Location/Qualifiers
1..8598
/organism="unknown"

source 1..177
/organism="unknown"
BASE COUNT 39 a 50 c 48 g 40 t
ORIGIN
Alignment Scores:
Pred. No.: 2.14e-28 Length: 177
Score: 323.00 Matches: 56
Percent Similarity: 100.00% Conservative: 1
Best Local Similarity: 98.25% Mismatches: 0
Query Match: 98.78% Indels: 0
DB: 6 Gaps: 0
09-833799-13B (1-57) x I96048 (1-177)
QY 1 AlaGlnGluProValLysGlyProValSerThrLysProGlySerCysProIleLeu 20
Db 1 GCACAGGAACCAAGTTAAAGGTCGGGTTCGACCAACCGGGCTCTTGGCCGATTATCTG 60
QY 21 IleArgCysAlaMetLeuAsnProProAsnArgCysLeuLysAspThrAspCysProGly 40
Db 61 ATCCGCTGCGTATCTGAACCGCGCAACCGTGTGTCTGAAAGACACTGACTGCCCGGGT 120
QY 41 IleLysLysCysCysGluGlySerCysGlyMetAlaCysPheValProGln 57
Db 121 ATCAAAAAATGCTGCGAAGGTTCTTGGCGTATGGCATGCTTCGTTCCGCAG 171
RESULT 20
LOCUS I96049 177 bp DNA linear PAT 01-DEC-1998
DEFINITION Sequence 7 from patent US 5734014.
ACCESSION I96049
VERSION I96049.1 GI:3940519
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 177)
AUTHORS Ishima,Y., Okawa,N., Yoshida,M., Amagaya,S. and Kaj,i,A.
TITLE Elafin derivative
JOURNAL Patent: US 5734014-A 7 31-MAR-1998;
FEATURES Location/Qualifiers
1..177
/organism="unknown"

source 1..177
/organism="unknown"
BASE COUNT 40 a 49 c 47 g 41 t
ORIGIN
Alignment Scores:
Pred. No.: 2.14e-28 Length: 177
Score: 323.00 Matches: 56
Percent Similarity: 100.00% Conservative: 1
Best Local Similarity: 98.25% Mismatches: 0
Query Match: 98.78% Indels: 0
DB: 6 Gaps: 0
09-833799-13B (1-57) x I96049 (1-177)
QY 1 AlaGlnGluProValLysGlyProValSerThrLysProGlySerCysProIleLeu 20
Db 1 GCACAGGAACCAAGTTAAAGGTCGGGTTCGACCAACCGGGCTCTTGGCCGATTATCTG 60
QY 21 IleArgCysAlaMetLeuAsnProProAsnArgCysLeuLysAspThrAspCysProGly 40
Db 61 ATCCGCTGCGTATCTGAACCGCGCAACCGTGTGTCTGAAAGACACTGACTGCCCGGGT 120
QY 41 IleLysLysCysCysGluGlySerCysGlyMetAlaCysPheValProGln 57
Db 121 ATCAAAAAATGCTGCGAAGGTTCTTGGCGTATGGCATGCTTCGTTCCGCAG 171
RESULT 21
LOCUS I74313 321 bp DNA linear PAT 03-APR-1998
DEFINITION Sequence 11 from patent US 5688641.

```

7685029
4 (bases 1 to 573)
Tamechika,I., Itakura,M., Saruta,Y., Furukawa,M., Kato,A.,
Tachibana,S. and Hirose,S.
Accelerated evolution in inhibitor domains of porcine elafin family
members
J. Biol. Chem. 271 (12), 7012-7018 (1996)
MEDLINE 96215132
PUBMED 8636131
FEATURES
    source          Location/Qualifiers
    1..573
        /organism="Bos taurus"
        /db_xref="taxon:9913"
        /tissue_type="tongue"
    1..405
        /gene="bTrappin-2"
    <1..405
        /gene="bTrappin-2"
        /function="transglutaminase substrate, putative proteinase
inhibitor"
        /codon_start=1
        /protein_id="CAA11184.1"
        /db_xref="GI:2764786"
        /db_xref="SPTREMBL:O46625"
        /translation="QEPVKGDPPVKGDPPVKGDPPVKQDPVKDQNPVRGQEPVKGD
PVKQDPPVKGDPPVKQEPVKQDPVKGDPPVKQDPVKGRIGGPLLTKPGSCPVLIRCAM
MNPENRCLRDAQCPGVKKCCGSGKTCMDPQ"
    <1..402
        /gene="bTrappin-2"
        /product="unnamed"
        /function="transglutaminase substrate, putative proteinase
inhibitor"
    mat_peptide     163 a 158 c 145 g 107 t

BASE COUNT      163 a 158 c 145 g 107 t
ORIGIN

Alignment Scores:
Pred. No.:      1.45e-19      Length:      573
Score:          250.00      Matches:    42
Percent Similarity: 77.42%      Conservative: 6
Best Local Similarity: 67.74%      Mismatches:  8
Query Match:    76.45%      Indels:      6
DB:              4      Gaps:         1

09-833799-13B (1-57) x BOT223216 (1-573)
QY  2  GlnGluProValLys-----GlyProValSerThrLysProGlySer 15
    |||:::|||||
Db 217 CAAGATCCAGTCAAAAGGCAAGGCCGAATCGGAGTCCACTCCTCACTAAGCCTGGGTCC 276

QY 16  CysProIleLeuIleArgCysAlaMetLeuAsnProProAsnArgCysLeuLysAsp 35
    |||||:::|||||
Db 277 TGCCCCAGGGTCTGATCCGGTGTGCCATGATGAACCCCTAACCGATGCTGAGGGAT 336

QY 36  ThrAspCysProGlyIleLysLysCysCysGluGlySerCysGlyMetAlaCysPheVal 55
    |||||:::|||||
Db 337 GCTCAGTCCCAGGGGTCAAGAAGTGTGTGAAGGCTCTTGTGGGAAGACCTGTATGGAT 396

QY 56  ProGln 57
    |||||
Db 397 CCCCAG 402

RESULT 23
PIGWAPA
LOCUS          PIGWAPA
DEFINITION    Pig DNA for elafin, complete cds.
ACCESSION     D50319
VERSION       D50319.1  GI:1228057
KEYWORDS      elafin.
SOURCE        Sus scrofa DNA, clone_lib:EMBL SP6/T7 clone:lambda WAP-1.
ORGANISM      Sus scrofa
              Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
              Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
REFERENCE     1 (sites)

```

AUTHORS Tamechika, I., Itakura, M., Saruta, Y., Furukawa, M., Kato, A.,
Tachibana, S. and Hirose, S.
TITLE Accelerated evolution in inhibitor domains of porcine elafin family
members
JOURNAL J. Biol. Chem. 271 (12), 7012-7018 (1996)
MEDLINE 96215132
REFERENCE 2 (bases 1 to 3693)
AUTHORS Hirose, S.
JOURNAL Unpublished
REFERENCE 3 (bases 1 to 3693)
AUTHORS Hirose, S.
TITLE Direct Submission
JOURNAL Submitted (18-APR-1995) Shigehisa Hirose, Tokyo Institute of
Technology, Department of Biological Sciences; 4259 Nagatsuta-cho,
Midori-ku, Yokohama, Kanagawa 226-8501, Japan
(E-mail: shirose@bio.titech.ac.jp, Tel: 045-924-5726,
Fax: 045-924-5824)

FEATURES

source
1. 3693
/organism="Sus scrofa"
/db_xref="taxon:9823"
/clone="lambda WAP-1"
/clone_lib="EMBL SP6/T7"
<1. .710
/number=1
<1. .628
join(629. .710, 1541. .1962)
/codon_start=1
/product="elafin"
/protein_id="BAA08854.1"
/db_xref="GI:1228058"
/translation="MRSRSFLVVLVFLVCGTLVAQAAGRIRPKGKGTKKILALVKG
QGVRGKQKVGQGVKQDGLGKSDPVKAQLPDKGDLGKSDSVKQDQPFKAQLPD
KLQDPVKAQPAIKRLILLTKPGSCPRILLRCLMVNPPNRLCLSDAQCPGLKKCEGFCG
KACMDPK"
629. .691
1777. .1959
/product="elafin"
711. .1540
/number=1
1541. .1963
/number=2
join(1963, 2445. .2601)
1964. .2444
/number=2
2141. .2402
/note="SINE (PRE-1)"
2445. .2601
/number=3
2580. .2585
polyA_signal
2601
polyA_site
repeat_region
3606. .3669
/note="TG repeat"

sig_peptide
mat_peptide

intron

exon

3'UTR

intron

repeat_unit

exon

polyA_signal

polyA_site

repeat_region

BASE COUNT 798 a 953 c 1112 g 830 t
ORIGIN

Alignment Scores:

Pred. No.: 1.78e-15 Length: 3693
Score: 221.50 Matches: 39
Percent Similarity: 69.84% Conservative: 5
Best Local Similarity: 61.90% Mismatches: 12
Query Match: 67.74% Indels: 7
DB: 4 Gaps: 1

09-833799-13B (1-57) x PIGWAPA (1-3693)

QY 2 GlnGluProValLysGlyPro-----ValSerThrLysProGly 14

Db 1771 CAAGATCCAGTGAAGGCCAACCTGCAATCAAAAGCTAATCTTACTCACCAGCCTGGC 1830

QY 15 SerCysProIleIleLeuIleArgCysAlaMetLeuAsnProProAsnArgCysLeuLys 34

Db 1831 TCCTGCCCCAGGATTCTGATCCGTTGCTTGATGGTCAATCCCCCTTAACAGGTGTTTGAGT 1890

QY 35 AspThrAspCysProGlyIleLysLysCysCysGluGlySerCysGlyMetAlaCysPhe 54
Db 1891 GATGCTCAGTGCCAGGCTCAAGAGTGCTGTGAAGGCTTTTTCGGGAGGCCTGTATG 1950
QY 55 ValProGln 57
Db 1951 GATCCCAAG 1959
RESULT 24
AB003281
LOCUS AB003281
DEFINITION Phacochoerus aethiopicus gene for elafin (trappin-2), partial cds.
ACCESSION AB003281
VERSION AB003281.1 GI:4887637
KEYWORDS elafin (trappin-2).
SOURCE Phacochoerus aethiopicus DNA.
ORGANISM Phacochoerus aethiopicus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Phacochoerus.
REFERENCE 1 (sites)
AUTHORS Furutani, Y., Kato, A., Yasue, H., Alexander, L.J., Beattie, C.W. and
Hirose, S.
TITLE Evolution of the trappin multigene family in the Suidae
JOURNAL J. Biochem. 124 (3), 491-502 (1998)
MEDLINE 98391820
REFERENCE 2 (bases 1 to 1128)
AUTHORS Hirose, S.
TITLE Direct Submission
JOURNAL Submitted (21-APR-1997) Shigehisa Hirose, Tokyo Institute of
Technology, Department of Biological Sciences; 4259 Nagatsuta-cho,
Midori-ku, Yokohama, Kanagawa 226-8501, Japan
(E-mail: shirose@bio.titech.ac.jp, Tel: 045-924-5726,
Fax: 045-924-5824)
FEATURES
source
1. 1128
/organism="Phacochoerus aethiopicus"
/db_xref="taxon:85517"
<211. .632
/codon_start=3
/product="elafin (trappin-2)"
/protein_id="BAA77825.1"
/db_xref="GI:4887638"
/translation="RPKGGTKRIHALVKQKQVRGEDQVKGVKQDGLGKQDQDPV
KAQLPDKGQDLGKGEDSVKQDPIKAQLPDQVQDPVKAQPAIKRLILLTKPGSCPRIL
IRCMVNPNNRCLSDAQCPGVKKCEGFCGKECLNPR"
211. .>632
exon
BASE COUNT 296 a 275 c 322 g 235 t
ORIGIN

Alignment Scores:
Pred. No.: 1.46e-15 Length: 1128
Score: 217.50 Matches: 38
Percent Similarity: 68.25% Conservative: 5
Best Local Similarity: 60.32% Mismatches: 13
Query Match: 66.51% Indels: 7
DB: 4 Gaps: 1

09-833799-13B (1-57) x AB003281 (1-1128)

QY 2 GlnGluProValLysGlyPro-----ValSerThrLysProGly 14

Db 441 CAAGATCCAGTCAAGGCCAACCTGCAATCAAAAGCTAATCTTACTTACCAAGCCTGGC 500

QY 15 SerCysProIleIleLeuIleArgCysAlaMetLeuAsnProProAsnArgCysLeuLys 34

Db 501 TCCTGCCCCAGGATTCTGATCCGTTGCTTGATGGTCAATCCCCCTTAACAGGTGTTTGAGT 560

QY 35 AspThrAspCysProGlyIleLysLysCysCysGluGlySerCysGlyMetAlaCysPhe 54

Db 561 GATGCTCAGTGCCAGGCTCAAGAGTGCTGTGAAGGCTTTTTCGGGAGGAATGTTTG 620

QY 55 ValProGln 57

Db 621 AATCCCAGG 629

|||||

RESULT 25
PIGWAPD
LOCUS PIGWAPD 464 bp mRNA linear MAM 14-APR-2000
DEFINITION Pig mRNA for elafin family member protein, complete cds.
ACCESSION D50322
VERSION D50322.1 GI:1228063
KEYWORDS elafin family member protein.
SOURCE Sus scrofa cDNA to mRNA, clone:WAP-5.
ORGANISM Sus scrofa
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
1 (sites)
Tamechika,I., Itakura,M., Saruta,Y., Furukawa,M., Kato,A.,
Tachibana,S. and Hirose,S.
Accelerated evolution in inhibitor domains of porcine elafin family
members
J. Biol. Chem. 271 (12), 7012-7018 (1996)
96215132
2 (sites)
Furutani,Y., Kato,A., Yasue,H., Alexander,L.J., Beattie,C.W. and
Hirose,S.
Evolution of the trappin multigene family in the Suidae
J. Biochem. 124 (3), 491-502 (1998)
98391820
3 (bases 1 to 464)
Hirose,S.
Direct Submission
Submitted (18-APR-1995) Shigehisa Hirose, Tokyo Institute of
Technology, Department of Biological Sciences; 4259 Nagatsuta-cho,
Midori-ku, Yokohama, Kanagawa 226-8501, Japan
(E-mail:shiroye@bio.titech.ac.jp, Tel:045-924-5726,
Fax:045-924-5824)
FEATURES
source Location/Qualifiers
1..464
/organism="Sus scrofa"
/db_xref="taxon:9823"
/clone="WAP-5"
<1..96
/number=1
1..19
/note="PCR primer"
<1..14
15..446
/codon_start=1
/product="elafin family member protein"
/protein_id="BAA08857.1"
/db_xref="GI:1228064"
translation="MRSRSFLVVLVVFLLICGLTVQAGRIIRPKGKTKTLALVKG
QGPVRGKQVKGQGVKQDLGKSDPVPKQLPDKGDPVKAQPAIKRLILLTKPGSC
PRILIRCLMVNPPNRCLSDAQCPGVKKCCGFCGKDCMDPK"
15..77
261..443
/product="unnamed"
97..447
/number=2
447..464
448..464
/number=3
complement(448..464)
/note="PCR primer"
BASE COUNT 126 a 111 c 124 g 103 t
ORIGIN

Alignment Scores:
Pred. No.: 7.48e-16 Length: 464
Score: 216.50 Matches: 38
Percent Similarity: 68.25% Conservative: 5
Best Local Similarity: 60.32% Mismatches: 13
Query Match: 66.21% Indels: 7
DB: 4 Gaps: 1

09-833799-13B (1-57) x PIGWAPD (1-464)

QY 2 GlnGluProValLysGlyPro-----ValSerThrLysProGly 14
|||||
Db 255 CAAGATCCAGTGAAAGGCCAACCTGCAATCAAACGCTAATCTTACTCACCAGCCTGGC 314
|||||
QY 15 SerCysProIleIleLeuIleArgCysAlaMetLeuAsnProProAsnArgCysLeuLys 34
|||||
Db 315 TCCTGCCCCAGGATCTGATCCGTTGCTTGATGCTCAATCCCCCTAACAGGTGTTGAGT 374
|||||
QY 35 AspThrAspCysProGlyIleLysLysCysCysGluGlySerCysGlyMetAlaCysPhe 54
|||||
Db 375 GATGCTCAGTGCCCGAGGGTCAAGAGTCTGTGAAGGCTTTTGCCTGGGAGGACTGTATG 434
|||||
QY 55 ValProGln 57
|||||
Db 435 GATCCCAAG 443
|||||

RESULT 26
D83668
LOCUS D83668 1034 bp DNA linear MAM 14-APR-2000
DEFINITION Sus scrofa gene for elafin homolog, exon2, partial cds.
ACCESSION D83668
VERSION D83668.1 GI:2055271
KEYWORDS elafin homolog; elastase inhibitor.
SOURCE Sus scrofa DNA, clone_lib:PCR product.
ORGANISM Sus scrofa
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
1 (bases 1 to 1034)
Tamechika,I., Itakura,M., Saruta,Y., Furukawa,M., Kato,A.,
Tachibana,S. and Hirose,S.
Accelerated evolution in the inhibitor domain of porcine elafin
family members
Journal of Biological Chemistry (1996) In press
2 (sites)
Furutani,Y., Kato,A., Yasue,H., Alexander,L.J., Beattie,C.W. and
Hirose,S.
Evolution of the trappin multigene family in the Suidae
J. Biochem. 124 (3), 491-502 (1998)
98391820
9722657
3 (bases 1 to 1034)
Hirose,S.
Direct Submission
Submitted (23-FEB-1996) Shigehisa Hirose, Tokyo Institute of
Technology, Department of Biological Sciences; 4259 Nagatsuta-cho,
Midori-ku, Yokohama, Kanagawa 226-8501, Japan
(E-mail:shiroye@bio.titech.ac.jp, Tel:045-924-5726,
Fax:045-924-5824)
On Apr 29, 1997 this sequence version replaced gi:1209380.
Sequence updated (21-Apr-1997).
FEATURES
source Location/Qualifiers
1..1034
/organism="Sus scrofa"
/db_xref="taxon:9823"
/clone_lib="PCR product"
<1..208
/number=1
1..24
/note="PCR primer"
209..559
/number=2
<209..558
/note="exon2 part"
/codon_start=3
/product="elafin homolog"
/protein_id="BAA12038.1"
/db_xref="GI:1209381"
translation="RPKKGKTKTLALVKQGPVRGKQVKGQDLGKSDPVPKQLPDKGDPVKAQPAIKRLILLTKPGSCPRILIRCLMVNPPNRCLSDAQCPGVKKCCGFCGKDCMDPK"
introns
primer_bind
exon
CDS

```
mat_peptide 373..555
/product="elafin homolog"
/function="elastase specific inhibitor"
/note="Mature protein has a characteristic
disulfide-linked structure called four disulfide core or
WAP motif"
560..>1034
/number=2
primer_bind 1010..>1034
/note="PCR primer"
BASE COUNT 263 a 262 c 298 g 211 t
ORIGIN
intron
560..>1034
/number=2
primer_bind 1010..>1034
/note="PCR primer"
BASE COUNT 263 a 262 c 298 g 211 t
ORIGIN
Alignment Scores:
Pred. No.: 1.73e-15 Length: 1034
Score: 216.50 Matches: 38
Percent Similarity: 68.25% Conservative: 5
Best Local Similarity: 60.32% Mismatches: 13
Query Match: 66.21% Indels: 7
DB: 4 Gaps: 1
09-833799-13B (1-57) x D83668 (1-1034)
QY 2 GlnGluProValLysGlyPro-----ValSerThrLysProGly 14
|||:::|||||
Db 367 CAAGATCCAGTGAAGCCCAACCTGCAATCAACGCTCTAATCTTACTCACCAAGCTGGC 426
|||:::|||||
QY 15 SerCysProIleLeuIleArgCysAlaMetLeuAsnProProAsnArgCysLeuLys 34
|||||
Db 427 TCCTGCCCCAGGATTCTGATCCGTTGCTTGATGCTCAATCCCTTAACAGGTGTTGAGT 486
|||||
QY 35 AspThrAspCysProGlyIleLysLysCysCysGluGlySerCysGlyMetAlaCysPhe 54
|||
Db 487 GATGCTCAGTGCAGGGGTCAAGAAAGTGTGTGAAGGCTTTTGCGGGAAGGACTGTATG 546
|||
QY 55 ValProGln 57
|||
Db 547 GATCCCAAG 555
RESULT 27
AB003283
LOCUS AB003283 1169 bp DNA linear MAM 26-MAY-1999
DEFINITION Pecari tajacu gene for trappin, partial cds.
ACCESSION AB003283
VERSION AB003283.1 GI:4887641
KEYWORDS trappin.
SOURCE Pecari tajacu DNA.
ORGANISM Pecari tajacu
REFERENCE 1 (sites)
AUTHORS Furutani,Y., Kato,A., Yasue,H., Alexander,L.J., Beattie,C.W. and Hirose,S.
TITLE Evolution of the trappin multigene family in the Suidae
JOURNAL J. Biochem. 124 (3), 491-502 (1998)
MEDLINE 98391820
REFERENCE 2 (bases 1 to 1169)
AUTHORS Hirose,S.
TITLE Direct Submission
JOURNAL Submitted (21-APR-1997) Shigehisa Hirose, Tokyo Institute of Technology, Department of Biological Sciences, 4259 Nagatsuta-cho, Midori-ku, Yokohama, Kanagawa 226-8501, Japan
(E-mail:shirose@bio.titech.ac.jp, Tel:045-924-5726, Fax:045-924-5824)
FEATURES
source
1..1169
/organism="Pecari tajacu"
/db_xref="taxon:9829"
<212..675
/codon_start=3
/product="trappin"
/protein_id="BAA77827.1"
/db_xref="GI:4887642"
/translation="RPKGQGTKKGHALIKGQDPVRGQGVKQDLPVKQALPDKGQDLV
KGQDPVKQDLPVKQDLPVKQALPDKGQDLVKGQDPVKQDLPVKQDLPVKQALPDKGQDLV
LPRRKPGFCPMIKIRCALFNPENRCLTDAGCPGARKCCIGSCGKACLNPRV"
212..675
exon 305 a 289 c 330 g 245 t
BASE COUNT
ORIGIN
Alignment Scores:
Pred. No.: 1.14e-13 Length: 1169
Score: 201.00 Matches: 35
Percent Similarity: 65.08% Conservative: 6
Best Local Similarity: 55.56% Mismatches: 14
Query Match: 61.47% Indels: 8
DB: 4 Gaps: 1
09-833799-13B (1-57) x AB003283 (1-1169)
QY 2 GlnGluProValLysGlyProValSerThr-----LysPro 13
|||:::|||||
Db 478 CAAGATCCAGTCAAGCCGAACTTGCACTGACAGCTGCTGCTTACCCCGCGTAAGCCT 537
|||:::|||||
QY 14 GlySerCysProIleLeuIleArgCysAlaMetLeuAsnProProAsnArgCysLeu 33
|||||
Db 538 GGCTTCTGCCCCATGATTAGATCCGTTGTGCCCTGTTCATCTCTTAACAGGTGTTG 597
|||||
QY 34 LysAspThrAspCysProGlyIleLysLysCysCysGluGlySerCysGlyMetAlaCys 53
|||||
Db 598 ACCGATGCTGGTGGTCCCGAGGAGGAGTGTGCTGCATAGGCTCTTGCGGGAAGGCTGT 657
|||||
QY 54 PheValPro 56
|||
Db 658 TTGAATCCT 666
RESULT 28
MAM223215
LOCUS MAM223215 270 bp mRNA linear PRI 15-DEC-1998
DEFINITION Macaca mulatta mRNA for putative strappin-2 protein, partial.
ACCESSION AJ223215
VERSION AJ223215.1 GI:2764783
KEYWORDS proteinase inhibitor; putative; strappin-2 gene; transglutaminase substrate.
SOURCE rhesus monkey.
ORGANISM Macaca mulatta
REFERENCE 1 (bases 1 to 270)
AUTHORS Zeeuwen,P.L., Hendriks,W., de Jong,W.W. and Schalkwijk,J.
TITLE Direct Submission
JOURNAL Submitted (06-JAN-1998) Zeeuwen P.L.J.M., Dermatology, University Hospital Nijmegen, PO BOX 9101, 6500 HB Nijmegen, THE NETHERLANDS
REFERENCE 2 (bases 1 to 270)
AUTHORS Zeeuwen,P.L., Hendriks,W., de Jong,W.W. and Schalkwijk,J.
TITLE Identification and sequence analysis of two new members of the SKALP/elafin and SPAI-2 gene family. Biochemical properties of the transglutaminase substrate motif and suggestions for a new nomenclature
JOURNAL J. Biol. Chem. 272 (33), 20471-20478 (1997)
MEDLINE 97400522
PUBMED 9252357
REFERENCE 3 (bases 1 to 270)
AUTHORS Tamechika,I., Itakura,M., Saruta,Y., Furukawa,M., Kato,A., Tachibana,S. and Hirose,S.
TITLE Accelerated evolution in inhibitor domains of porcine elafin family members
JOURNAL J. Biol. Chem. 271 (12), 7012-7018 (1996)
MEDLINE 96215132
PUBMED 8636131
REFERENCE 4 (bases 1 to 270)
AUTHORS Molhuizen,H.O., Alkemade,H.A., Zeeuwen,P.L., de Jongh,G.J., Wieringa,B. and Schalkwijk,J.
TITLE SKALP/elafin: an elastase inhibitor from cultured human keratinocytes. Purification, cDNA sequence, and evidence for
```

JOURNAL transglutaminase cross-linking
MEDLINE J. Biol. Chem. 268 (16), 12028-12032 (1993)
PUBMED 93280175
7685029

FEATURES
source Location/Qualifiers
1..270
/organism="Macaca mulatta"
/db_xref="taxon:9544"
/tissue_type="skin"
1..270
/gene="sTrappin-2"
1..270
/gene="sTrappin-2"
1..270
/function="transglutaminase substrate, putative proteinase inhibitor"
/codon_start=2
/protein_id="CAA11183.1"
/db_xref="GI:2764784"
/db_xref="SPTREMBL:O46643"
/translation="VVFLIAGMLVVEAAVTGVPVKGDQTVKGRVFPNGQDPVKGVQS
VKQDRVKGKRGVPGVSTKPGSCPILIRCAMLNPPNRLKDTD"
1..43
/gene="sTrappin-2"
41..270
/gene="sTrappin-2"
/product="unnamed"
/function="transglutaminase substrate, putative proteinase inhibitor"

BASE COUNT 64 a 66 c 74 g 66 t
ORIGIN

sig_peptide
mat_peptide

Alignment Scores:
Pred. No.: 3.53e-12 Length: 270
Score: 182.00 Matches: 33
Percent Similarity: 97.06% Conservative: 0
Best Local Similarity: 97.06% Mismatches: 1
Query Match: 55.66% Indels: 0
DB: 9 Gaps: 0

09-833799-13B (1-57) x MAM223215 (1-270)

QY 4 ProValLysGlyProValSerThrLysProGlySerCysProIleIleLeuIleArgCys 23
|||||
Db 167 CCAGTCAAAGGTCCTCCACTAAGCCTGGCTCCTGCCCAATATCTTGATCCGGTGC 226
|||||

QY 24 AlaMetLeuAsnProAsnArgCysLeuLysAspThrAsp 37
|||||
Db 227 GCCATGTTGAATCCCTAACCCTGCTTGAAGATACCTGAC 268
|||||

RESULT 29
AB003282
LOCUS 1325 bp DNA linear MAM 26-MAY-1999
DEFINITION Phacochoerus aethiopicus gene for SPAI (trappin-1), partial cds.
ACCESSION AB003282
VERSION AB003282.1 GI:4887639
KEYWORDS SPAI (trappin-1).
SOURCE Phacochoerus aethiopicus
ORGANISM Phacochoerus aethiopicus
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Phacochoerus.
1 (sites)
AUTHORS Furutani,Y., Kato,A., Yasue,H., Alexander,L.J., Beattie,C.W. and Hirose,S.
TITLE Evolution of the trappin multigene family in the Suidae
JOURNAL J. Biochem. 124 (3), 491-502 (1998)
MEDLINE 98391820
REFERENCE 2 (sites)
AUTHORS Furutani,Y., Kato,A., Yasue,H. and Hirose,S.
TITLE (Very Recent) evolution in suidae of the trappin multigene family with unusually conserved intron sequences
JOURNAL Unpublished
REFERENCE 3 (bases 1 to 1325)
AUTHORS Hirose,S.

TITLE Direct Submission
JOURNAL Submitted (21-APR-1997) Shigehisa Hirose, Tokyo Institute of Technology, Department of Biological Sciences; 4259 Nagatsuta-cho, Midori-ku, Yokohama, Kanagawa 226-8501, Japan
(E-mail:shirose@bio.titech.ac.jp, Tel:045-924-5726, Fax:045-924-5824)

FEATURES
source Location/Qualifiers
1..1325
/organism="Phacochoerus aethiopicus"
/db_xref="taxon:85517"
1..828
/codon_start=3
/product="SPAI (trappin-1)"
/protein_id="BAA77826.1"
/db_xref="GI:4887640"
/translation="GPKSQGQDPVEGQDHDEGQGPVKVEILDIGQDLVKGQDPVKGD
LVKSQDPVKAELPDIGQDVVKGQDPVKQDLVKSQDPVTAELPDIGQDVVKGQDPVKG
QDLVKSQDPVTAELPDIGQDVVKGHEPVEGQDPINAIQIPDKVQDPIKAQPAVQGLLFL
SKRGRCPWILLRCLANPSNKCWRDYDCPGVKKCEGFCGKDCLYPK"
203..828
exon 363 a 319 c 367 g 276 t
ORIGIN

Alignment Scores:
Pred. No.: 2.14e-11 Length: 1325
Score: 181.50 Matches: 30
Percent Similarity: 63.49% Conservative: 10
Best Local Similarity: 47.62% Mismatches: 16
Query Match: 55.50% Indels: 7
DB: 4 Gaps: 1

09-833799-13B (1-57) x AB003282 (1-1325)

QY 2 GlnGluProValLysGlyProValSer-----ThrLysProGly 14
|||||
Db 637 CAAGATCCAATCAAGCCCAACCTGCAGTCCAGGTCTACTCTTCTCTAAGCGTGGC 696
|||||

QY 15 SerCysProIleIleLeuIleArgCysAlaMetLeuAsnProProAsnArgCysLeuLys 34
|||||
Db 697 CGTGCCTTGGATTCTTCCGTTGCCCGCTGGCCATCCCTCTAACCAAGTTGGAGA 756
|||||

QY 35 AspThrAspCysProGlyIleLysLysCysCysGluGlySerCysGlyMetAlaCysPhe 54
|||||
Db 757 GATTATGACTGTCCAGGGGTCAAGAAAGTGTCTGAAGGCTTTTGGCGGAAGGATTGTTG 816
|||||

QY 55 ValProGln 57
|||||
Db 817 TATCCCAAG 825
|||||

RESULT 30
PIGSPA12S2
LOCUS 483 bp DNA linear MAM 14-APR-2000
DEFINITION Porcine DNA for SPAI-2, exon 2.
ACCESSION D17755
VERSION D17755.1 GI:1054614
KEYWORDS SPAI-2.
SEGMENT 2 of 3
SOURCE Sus scrofa cell_line NMS38 (library: lambda EMBL3 SP6/T7) DNA, clone lamdapSG1.
ORGANISM Sus scrofa
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
1 (bases 1 to 483)
AUTHORS Kuroki,J., Hosoya,T., Itakura,M., Hirose,S., Tamechika,I., Yoshimoto,T., Ghoneim,M.A., Nara,K., Kato,A., Suzuki,Y., Furukawa,M. and Tachibana,S.
TITLE Cloning, characterization, and tissue distribution of porcine SPAI, a protein with a transglutaminase substrate domain and the WAP motif
JOURNAL J. Biol. Chem. 270 (38), 22428-22433 (1995)
MEDLINE 95403443
REFERENCE 2 (bases 1 to 483)
AUTHORS Hirose,S.

REFERENCE 1 (sites)
AUTHORS Kuroki,J., Hosoya,T., Itakura,M., Hirose,S., Tamechika,I., Yoshimoto,T., Ghoneim,M.A., Nara,K., Kato,A., Suzuki,Y., Furukawa,M. and Tachibana,S.
TITLE Cloning, characterization, and tissue distribution of porcine SPAl, a protein with a transglutaminase substrate domain and the WAP motif
JOURNAL J. Biol. Chem. 270 (38), 22428-22433 (1995)
MEDLINE 95403443
REFERENCE 2 (bases 1 to 789)
AUTHORS Hirose,S.
TITLE Direct Submission
JOURNAL Submitted (23-FEB-1996) Shigehisa Hirose, Tokyo Institute of Technology, Department of Biological Sciences; Ookayama, Meguro-ku, Tokyo 152, Japan (Tel:81-3-3726-1111, Fax:81-3-3729-0335)
COMMENT On or before Mar 17, 1999 this sequence version replaced gi:2078450, gi:1054611, gi:2077950.
D17753:Submitted (27-Sep-1993) to DDBJ by:Shigehisa Hirose.
FEATURES
source Location/Qualifiers
1. .789
/organism="Sus scrofa"
/db_xref="taxon:9823"
/cell_line="NM514"
/clone_lib="lambda gt10"
61. .624
/codon_start=1
/product="proSPAI-2"
/protein_id="BAA12037.1"
/db_xref="GI:1304181"
/translation="MRSRSFLVLVAVFLICETLVAQRDLIRGPKGQGQDPVEGQDQD EGPGPVKVEILDIGQDPVKGDVPKGDVPKGDVPKGDVPKQDLPVKSDPVKAELPDIGQD VVKGHEPVEGQDPPVNAQLPKVQDPPVKAQPAVPGRFLLSKRGHCPRILFRCLSNPSN KCWRDYDCPGVKKCCGFCGKDCLYPK"
61. .123
124. .621
/product="proSPAI-2"
439. .621
/product="SPAI-2"
722
polyA_site
polyA_signal 761. .766
BASE COUNT 201 a 215 c 197 g 176 t
ORIGIN
Alignment Scores:
Pred. No.: 2.09e-11 Length: 789
Score: 179.50 Matches: 31
Percent Similarity: 61.90% Conservative: 8
Best Local Similarity: 49.21% Mismatches: 17
Query Match: 54.89% Indels: 7
DB: 4 Gaps: 1
09-833799-13B (1-57) x DMY245 (1-789)
QY 2 GlnGluProValLysGlyPro-----ValSerThrLysProGly 14
Db 433 CAAGATCCAGTCAAGGCCAACCTGCAGTCCCAGGTCGATTCTCTCTAAGCGTGGC 492
QY 15 SerCysProIleIleLeuIleArgCysAlaMetLeuAsnProProAsnArgCysLeuLys 34
Db 493 CACTGCCCTAGGATTCTTTTCGTTGCCGCTGAGCAATCCCTCTAACAAGTGTGGAGA 552
QY 35 AspThrAspCysProGlyIleLysLysCysCysGluGlySerCysGlyMetAlaCysPhe 54
Db 553 GATTATGACTGTCCAGGGGTCAAGAAGTGTCTGAAGGCTTTTGGCGAAGGATTGTTG 612
QY 55 ValProGln 57
Db 613 TATCCCAAG 621
RESULT 33
PIGWAPB PIGWAPB 3782 bp DNA linear MAM 14-APR-2000
LOCUS PIGWAPB 3782 bp DNA for SPAl-2, complete cds.
DEFINITION Pig DNA for SPAl-2, complete cds.

ACCESSION D50320
VERSION D50320.1 GI:1228059
KEYWORDS SPAl-2.
SOURCE Sus scrofa DNA, clone_lib:EMBL SP6/T7 clone:lambda WAP-2.
ORGANISM Sus scrofa
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
REFERENCE 1 (sites)
AUTHORS Tamechika,I., Itakura,M., Saruta,Y., Furukawa,M., Kato,A., Tachibana,S. and Hirose,S.
TITLE Accelerated evolution in inhibitor domains of porcine elafin family members
JOURNAL J. Biol. Chem. 271 (12), 7012-7018 (1996)
MEDLINE 96215132
REFERENCE 2 (bases 1 to 3782)
AUTHORS Hirose,S.
JOURNAL Unpublished
REFERENCE 3 (bases 1 to 3782)
AUTHORS Hirose,S.
TITLE Direct Submission
JOURNAL Submitted (18-APR-1995) Shigehisa Hirose, Tokyo Institute of Technology, Department of Biological Sciences; 4259 Nagatsuta-cho, Midori-ku, Yokohama, Kanagawa 226-8501, Japan (E-mail:shihiro@bio.titech.ac.jp, Tel:045-924-5726, Fax:045-924-5824)
FEATURES
source Location/Qualifiers
1. .3782
/organism="Sus scrofa"
/db_xref="taxon:9823"
/clone_lib="lambda WAP-2"
/clone_lib="EMBL SP6/T7"
1. .710
/number=1
/evidence=experimental
1. .628
join(629. .710,1552. .2033)
/codon_start=1
/evidence=experimental
/product="SPAI-2"
/protein_id="BAA08855.1"
/db_xref="GI:1228060"
/translation="MRSRSFLVLVAVFLICETLVAQRDLIRGPKGQGQDPVEGQDQD EGPGPVKVEILDIGQDPVKGDVPKGDVPKGDVPKQDLPVKSDPVKAELPDIGQD VVKGHEPVEGQDPPVNAQLPKVQDPPVKAQPAVPGRFLLSKRGHCPRILFRCLSNPSN KCWRDYDCPGVKKCCGFCGKDCLYPK"
629. .691
/evidence=experimental
1848. .2030
/product="SPAI-2"
/evidence=experimental
711. .1551
/number=1
/evidence=experimental
1552. .2034
/number=2
/evidence=experimental
join(2034,2545. .2701)
2035. .2544
/number=2
/evidence=experimental
2240. .2502
/note="SINE (PRE-1)"
2545. .2701
/number=3
/evidence=experimental
2680. .2685
2701
/evidence=experimental
3671. .3742
/note="TG repeat"
BASE COUNT 823 a 972 c 1149 g 838 t
ORIGIN

```
Alignment Scores:
Pred. No.:      1.08e-10      Length:      3782
Score:          179.50      Matches:      31
Percent Similarity: 61.90%      Conservative: 8
Best Local Similarity: 49.21%      Mismatches: 17
Query Match:    54.89%      Indels:    7
DB:             4      Gaps:      1

09-833799-13B (1-57) x PIGWAPB (1-3782)
QY      2  GlnGluProValLysGlyPro-----ValSerThrLysProGly 14
      |||:::|||||
Db      1842  CAAGATCCAGTCAAGCCCAACCTGCAGTCCAGGTCGATTCCTTCTCTAAGCGTGGC 1901
      |||:::|||||
QY      15  SerCysProIleIleLeuIleArgCysAlaMetLeuAsnProProAsnArgCysLeuLys 34
      |||:::|||||
Db      1902  CACTGCCCTAGGATTCTTTTTCGTTGCCCGCTGAGCAATCCCTCTAACAAGTGTGGAGA 1961
      |||:::|||||
QY      35  AspThrAspCysProGlyIleLysLysCysCysGluGlySerCysGlyMetAlaCysPhe 54
      |||:::|||||
Db      1962  GATTATGACTGTCCAGGGGTCAAGAAGTGTGTGAAGGCTTTTTCGGGAAGGATTGTTTG 2021
      |||:::|||||
QY      55  ValProGln 57
      |||:::
Db      2022  TATCCCAAG 2030

RESULT 34
PIGWAPE
LOCUS      PIGWAPB      578 bp      mRNA      linear      MAM 14-APR-2000
DEFINITION  Pig mRNA for elafin family member protein, complete cds.
ACCESSION  D50323
VERSION    D50323.1 GI:1228065
KEYWORDS   elafin family member protein.
SOURCE     Sus scrofa cDNA to mRNA, clone:WAP-4.
ORGANISM   Sus scrofa
REFERENCE  Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS    Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
TITLE      Tamechika, I., Itakura, M., Saruta, Y., Furukawa, M., Kato, A.,
            Tachibana, S. and Hirose, S.
            Accelerated evolution in inhibitor domains of porcine elafin family
            members
JOURNAL    J. Biol. Chem. 271 (12), 7012-7018 (1996)
MEDLINE    96215132
REFERENCE  2 (sites)
AUTHORS    Furutani, Y., Kato, A., Yasue, H., Alexander, L.J., Beattie, C.W. and
            Hirose, S.
TITLE      Evolution of the trappin multigene family in the Suidae
JOURNAL    J. Biochem. 124 (3), 491-502 (1998)
MEDLINE    98391820
REFERENCE  3 (bases 1 to 578)
AUTHORS    Hirose, S.
TITLE      Direct Submission
JOURNAL    Submitted (18-APR-1995) Shigehisa Hirose, Tokyo Institute of
            Technology, Department of Biological Sciences, 4259 Nagatsuta-cho,
            Midori-ku, Yokohama, Kanagawa 226-8501, Japan
            (E-mail: shirose@bio.titech.ac.jp, Tel:045-924-5726,
            Fax:045-924-5824)
FEATURES   Location/Qualifiers
            source      1..578
                        /organism="Sus scrofa"
                        /db_xref="taxon:9823"
                        /clone="WAP-4"
                        <1..96
                        /number=1
            primer_bind  1..19
                        /note="PCR primer"
                        <1..14
                        15..560
                        /codon_start=1
                        /product="elafin family member protein"
                        /protein_id="BAA08858.1"
                        /db_xref="GI:1864016"

exon
primer_bind
5'UTR
CDS

/translation="MRSRSFLVLVAVFLICETLVAQRPEKIRGPKGQGPVEGQDQD
KGQGPVKVEILDIGQDLVKRDPVKGDPPVKGQDLVKSQDPVKAELPDIGQDVVKGHD
PVEGQDPVNAQLPDKVQDPVKAQPPVQGRLLHYKPGLCPLWIFLRCPLPKPPNKCWRDS
HCPGVMKCEGFCGNECSYPR"
15..77
375..557
/product="unnamed"
exon
97..561
/number=2
561..578
562..578
/number=3
primer_bind
complement(562..578)
/note="PCR primer"
BASE COUNT 169 a 141 c 138 g 130 t
ORIGIN
Alignment Scores:
Pred. No.:      2.55e-11      Length:      578
Score:          177.50      Matches:      30
Percent Similarity: 67.24%      Conservative: 9
Best Local Similarity: 51.72%      Mismatches: 18
Query Match:    54.28%      Indels:    1
DB:             4      Gaps:      1

09-833799-13B (1-57) x PIGWAPE (1-578)
QY      1  AlaGlnGluProValLysGlyPro--ValSerThrLysProGlySerCysProIleIle 19
      |||:::|||||
Db      384  GCCCAACCTCCAGTCCCAAGTCGATTACTTCACTATAAGCCTGGCCTCTGCCCTTGGATT 443
      |||:::|||||
QY      20  LeuIleArgCysAlaMetLeuAsnProProAsnArgCysLeuLysAspThrAspCysPro 39
      |||:::|||||
Db      444  TTTTACGTTGCCCGCTGCCCAACCTCTCTAACAATGTTGGAGAGATTCTCACTGCCCA 503
      |||:::|||||
QY      40  GlyIleLysLysCysCysGluGlySerCysGlyMetAlaCysPheValProGln 57
      |||:::|||||
Db      504  GGGTCATGAAGTCTGTGAAGGCTTTTTCGGGGGAATGAATGTTCTGTATCCCCGA 557
      |||:::|||||

RESULT 35
AF178426
LOCUS      AF178426      490 bp      mRNA      linear      ROD 29-AUG-1999
DEFINITION  Rattus norvegicus secretory leukocyte protease inhibitor mRNA,
            complete cds.
ACCESSION  AF178426
VERSION    AF178426.1 GI:5802679
KEYWORDS
SOURCE     Rattus norvegicus.
ORGANISM   Rattus norvegicus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
            Rattus.
REFERENCE  1 (bases 1 to 490)
AUTHORS    Song, X., Zeng, L., Jin, W., Thompson, J., Mizel, D.E., Lei, K.,
            Billinghamurst, R.C., Poole, A.R. and Wahl, S.M.
TITLE      Secretory leukocyte protease inhibitor suppresses the inflammation
            and joint damage of bacterial cell wall-induced arthritis
JOURNAL    J. Exp. Med. 190 (4), 535-542 (1999)
MEDLINE    99380355
PUBMED     10449524
REFERENCE  2 (bases 1 to 490)
AUTHORS    Song, X.-Y. and Zeng, L.
TITLE      Direct Submission
JOURNAL    Submitted (17-AUG-1999) Oral Infection & Immunity Branch, National
            Institute of Dental & Craniofacial research, 9000 Rockville Pike,
            Bethesda, MD 20892
FEATURES   Location/Qualifiers
            source      1..490
                        /organism="Rattus norvegicus"
                        /strain="Lewis"
                        /db_xref="taxon:10116"
                        9..404
                        /note="slpi"
            CDS
```

```
/codon_start=1
/product="secretory leukocyte protease inhibitor"
/protein_id="AAD51758.1"
/db_xref="GI:5802680"
/translation="MKSCGLFPLMVLALGLVAPWSVEGGKNDAIKIGACPAKPKPAQC
LKLEKECGTDWECPGKQRCQDTCGFKCLNPVPIRGVKKKPGRCVKFQGKCLMLNP
PNKQNDGQCDGKYKCEGMCVKCLPPV"
BASE COUNT 118 a 121 c 131 g 120 t
ORIGIN

Alignment Scores:
Pred. No.: 1.17e-10 Length: 490
Score: 171.00 Matches: 28
Percent Similarity: 60.38% Conservative: 4
Best Local Similarity: 52.83% Mismatches: 21
Query Match: 52.29% Indels: 0
DB: 10 Gaps: 0

09-833799-13B (1-57) x AFI78426 (1-490)
QY 4 ProValLysGlyProValSerThrLysProGlySerCysProIleIleLeuIleArgCys 23
Db 240 CCCATTCGTGGACCAAGTGAAGAACGCTGGAGGTGCGTCAAAATTTCAAGGAAATGT 299

QY 24 AlaMetLeuAsnProAsnArgCysLeuLysAspThrAspCysProGlyIleLysLys 43
Db 300 CTGATGCTTAACCTCCCAATAAGTCCAGAAATGACGGCCAGTGTGATGGCAATACAAA 359

QY 44 CysCysGluGlySerCysGlyMetAlaCysPheValPro 56
Db 360 TGTGTGAGGGCATGTGTGGGAAAGTGTGCTTCCCCCA 398

RESULT 36
AB011010
LOCUS Bos taurus gene for Trappin-6, partial cds. 495 bp DNA linear MAM 14-APR-2000
DEFINITION Bos taurus gene for Trappin-6, partial cds.
ACCESSION AB011010
VERSION AB011010.1 GI:3132273
KEYWORDS Trappin-6.
SOURCE Bos taurus DNA.
ORGANISM Bos taurus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
Bovidae; Bovinae; Bos.
REFERENCE 1 (sites)
AUTHORS Furutani,Y., Kato,A., Yasue,H., Alexander,L.J., Beattie,C.W. and Hirose,S.
TITLE Evolution of the trappin multigene family in the Suidae
JOURNAL J. Biochem. 124 (3), 491-502 (1998)
MEDLINE 98391820
REFERENCE 2 (bases 1 to 495)
AUTHORS Hirose,S.
TITLE Direct Submission
JOURNAL Submitted (10-FEB-1998) Shigehisa Hirose, Tokyo Institute of Technology, Department of Biological Sciences; 4259 Nagatsuta-cho, Midori-ku, Yokohama, Kanagawa 226-8501, Japan
(E-mail:shirose@bio.titech.ac.jp, Tel:+81-45-924-5726, Fax:+81-45-924-5824)
FEATURES
source Location/Qualifiers
1..495
/organism="Bos taurus"
/db_xref="taxon:9913"
<1..36
/number=1
primer_bind 1..24
/note="PCR primer site"
<37..329
/codon_start=3
/product="Trappin-6"
/protein_id="BAA28148.1"
/db_xref="GI:3132274"
/translation="SPKGGNVVFNKGKPVNGQSPDKGQDPVKGDVPVKQDVVVAQD
RAGLPFKRGLCPVRVIRHCLNPNPNCWRDAHCPGAKKCEGCGKTCMNP"
/codon_start=1
/product="secretory leukocyte protease inhibitor"
/protein_id="AAD51758.1"
/db_xref="GI:5802680"
/translation="MKSCGLFPLMVLALGLVAPWSVEGGKNDAIKIGACPAKPKPAQC
LKLEKECGTDWECPGKQRCQDTCGFKCLNPVPIRGVKKKPGRCVKFQGKCLMLNP
PNKQNDGQCDGKYKCEGMCVKCLPPV"
BASE COUNT 118 a 121 c 131 g 120 t
ORIGIN

Alignment Scores:
Pred. No.: 1.17e-10 Length: 490
Score: 171.00 Matches: 28
Percent Similarity: 60.38% Conservative: 4
Best Local Similarity: 52.83% Mismatches: 21
Query Match: 52.29% Indels: 0
DB: 10 Gaps: 0

09-833799-13B (1-57) x AFI78426 (1-490)
QY 4 ProValLysGlyProValSerThrLysProGlySerCysProIleIleLeuIleArgCys 23
Db 240 CCCATTCGTGGACCAAGTGAAGAACGCTGGAGGTGCGTCAAAATTTCAAGGAAATGT 299

QY 24 AlaMetLeuAsnProAsnArgCysLeuLysAspThrAspCysProGlyIleLysLys 43
Db 300 CTGATGCTTAACCTCCCAATAAGTCCAGAAATGACGGCCAGTGTGATGGCAATACAAA 359

QY 44 CysCysGluGlySerCysGlyMetAlaCysPheValPro 56
Db 360 TGTGTGAGGGCATGTGTGGGAAAGTGTGCTTCCCCCA 398

RESULT 36
AB011010
LOCUS Bos taurus gene for Trappin-6, partial cds. 495 bp DNA linear MAM 14-APR-2000
DEFINITION Bos taurus gene for Trappin-6, partial cds.
ACCESSION AB011010
VERSION AB011010.1 GI:3132273
KEYWORDS Trappin-6.
SOURCE Bos taurus DNA.
ORGANISM Bos taurus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
Bovidae; Bovinae; Bos.
REFERENCE 1 (sites)
AUTHORS Furutani,Y., Kato,A., Yasue,H., Alexander,L.J., Beattie,C.W. and Hirose,S.
TITLE Evolution of the trappin multigene family in the Suidae
JOURNAL J. Biochem. 124 (3), 491-502 (1998)
MEDLINE 98391820
REFERENCE 2 (bases 1 to 495)
AUTHORS Hirose,S.
TITLE Direct Submission
JOURNAL Submitted (10-FEB-1998) Shigehisa Hirose, Tokyo Institute of Technology, Department of Biological Sciences; 4259 Nagatsuta-cho, Midori-ku, Yokohama, Kanagawa 226-8501, Japan
(E-mail:shirose@bio.titech.ac.jp, Tel:+81-45-924-5726, Fax:+81-45-924-5824)
FEATURES
source Location/Qualifiers
1..495
/organism="Bos taurus"
/db_xref="taxon:9913"
<1..36
/number=1
primer_bind 1..24
/note="PCR primer site"
<37..329
/codon_start=3
/product="Trappin-6"
/protein_id="BAA28148.1"
/db_xref="GI:3132274"
/translation="SPKGGNVVFNKGKPVNGQSPDKGQDPVKGDVPVKQDVVVAQD
RAGLPFKRGLCPVRVIRHCLNPNPNCWRDAHCPGAKKCEGCGKTCMNP"
/codon_start=1
/product="secretory leukocyte protease inhibitor"
/protein_id="AAD51758.1"
/db_xref="GI:5802680"
/translation="MKSCGLFPLMVLALGLVAPWSVEGGKNDAIKIGACPAKPKPAQC
LKLEKECGTDWECPGKQRCQDTCGFKCLNPVPIRGVKKKPGRCVKFQGKCLMLNP
PNKQNDGQCDGKYKCEGMCVKCLPPV"
BASE COUNT 118 a 121 c 131 g 120 t
ORIGIN
```

```
exon 37..329
primer_bind 472..495
/note="PCR primer site"
BASE COUNT 134 a 99 c 163 g 99 t
ORIGIN

Alignment Scores:
Pred. No.: 1.54e-10 Length: 495
Score: 170.00 Matches: 30
Percent Similarity: 56.06% Conservative: 7
Best Local Similarity: 45.45% Mismatches: 19
Query Match: 51.99% Indels: 10
DB: 4 Gaps: 1

09-833799-13B (1-57) x AB011010 (1-495)
QY 2 GlnGluProValLysGlyPro-----ValSerThr 11
Db 129 CAAGACCCAGTCAAGAGCAAGATGTAGTCGTAGCACAAAGACCGAGCCGACTTCCATTC 188

QY 12 LysProGlySerCysProIleIleLeuIleArgCysAlaMetLeuAsnProAsnArg 31
Db 189 AAGCGTGGCTTATGCCCCAGGGTTCGGATCCACTGCAACTTGTGGAATCCCCCTAACAG 248

QY 32 CysLeuLysAspThrAspCysProGlyIleLysLysCysCysGluGlySerCysGlyMet 51
Db 249 TGTGGAGAGATGCTCACTGCCAGGGGCGGCGGAGAGTGTGTGAAGGCTTTTGTGGGAAG 308

QY 52 AlaCysPheValProGln 57
Db 309 ACCTGTATGAATCCCCGG 326

RESULT 37
AB003284
LOCUS Hippopotamus amphibius gene for trappin, partial cds. 734 bp DNA linear MAM 26-MAY-1999
DEFINITION Hippopotamus amphibius gene for trappin, partial cds.
ACCESSION AB003284
VERSION AB003284.1 GI:4887643
KEYWORDS trappin.
SOURCE Hippopotamus amphibius DNA.
ORGANISM Hippopotamus amphibius
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Cetartiodactyla; Hippopotamidae; Hippopotamus.
REFERENCE 1 (sites)
AUTHORS Furutani,Y., Kato,A., Yasue,H., Alexander,L.J., Beattie,C.W. and Hirose,S.
TITLE Evolution of the trappin multigene family in the Suidae
JOURNAL J. Biochem. 124 (3), 491-502 (1998)
MEDLINE 98391820
REFERENCE 2 (bases 1 to 734)
AUTHORS Hirose,S.
TITLE Direct Submission
JOURNAL Submitted (21-APR-1997) Shigehisa Hirose, Tokyo Institute of Technology, Department of Biological Sciences; 4259 Nagatsuta-cho, Midori-ku, Yokohama, Kanagawa 226-8501, Japan
(E-mail:shirose@bio.titech.ac.jp, Tel:045-924-5726, Fax:045-924-5824)
FEATURES
source Location/Qualifiers
1..734
/organism="Hippopotamus amphibius"
/db_xref="taxon:9833"
<1..545
/codon_start=3
/product="trappin"
/protein_id="BAA77828.1"
/db_xref="GI:4887644"
/translation="HPKGGTKGNVLFKQDPPVKQDPVKGQYPVKGQDPVKGQDPV
KGQYPVKGQDPVKQDPPVKQDPPVKQDPPVKQDPPVKQDPPVKQDPPVKQD
PVKQDPPVKQYPVKQDPPVKQDPPVKQDPPVKQDPPVKQDPPVKQDPPVKQD
CPRNKNCPPASCCGKICLKFP"
BASE COUNT 245 a 158 c 198 g 133 t
ORIGIN
```

Alignment Scores:
Pred. No.: 6.64e-10 Length: 734
Score: 166.00 Matches: 28
Percent Similarity: 61.54% Conservative: 4
Best Local Similarity: 53.85% Mismatches: 20
Query Match: 50.76% Indels: 0
DB: 4 Gaps: 0

09-833799-13B (1-57) x AB003284 (1-734)

QY 2 GlnGluProValSerThrLysProGlySerCysProIleLeuIle 21
Db 375 CAAGATCCAGTCAAAGTCCCATTCCTTGCTGCGGAAGGTGCTGCCCAAGATTGGATT 434
QY 22 ArgCysAlaMetLeuAsnProProAsnArgCysLeuLysAspThrAspCysProGlyIle 41
Db 435 GAGTGCTCCACATTAAATCCCCAAACCGTGTGAGAGATGCTCAGTGCCCAAGGAAC 494
QY 42 LysLysCysCysGluGlySerCysGlyMetAlaCys 53
Db 495 AAGAATTGCTGTCGCGCCTCTTGGCGGAAGATCTGT 530

RESULT 38
PIGALP
LOCUS Porcine antileukoprotease mRNA, complete cds. MAM 07-MAR-1995
DEFINITION
ACCESSION M57446
VERSION M57446.1 GI:164319
KEYWORDS antileukoprotease.
SOURCE Porcine uterus. cDNA to mRNA, clone pALP.
ORGANISM Sus scrofa
REFERENCE Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
AUTHORS Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
TITLE 1 (bases 1 to 600)
Farmer,S.J., Fliss,A.E. and Simmen,R.C.
Complementary DNA cloning and regulation of expression of the
messenger RNA encoding a pregnancy-associated porcine uterine
protein related to human antileukoprotease
Mol. Endocrinol. 4 (8), 1095-1104 (1990)
JOURNAL
MEDLINE 91155942
PUBMED 2293019

FEATURES
source Location/Qualifiers
1..600
/organism="Sus scrofa"
/db_xref="taxon:9823"
/clone="pALP"
/sex="female"
/tissue_type="uterus"
1..600
/gene="ALP"
44..391
/gene="ALP"
/codon_start=1
/product="antileukoprotease"
/protein_id="AAA63446.1"
/db_xref="GI:164320"
KCCRDTCALKCLNPVAITNPVKPKGCPVYGCMMMLNPPNHCKTDSQCLGDLKCKK
SMCGKVLTPVKA"
polyA_site 600
BASE COUNT 146 a 156 c 151 g 147 t
ORIGIN

Alignment Scores:
Pred. No.: 6.98e-10 Length: 600
Score: 165.00 Matches: 25
Percent Similarity: 59.62% Conservative: 6
Best Local Similarity: 48.08% Mismatches: 21
Query Match: 50.46% Indels: 0
DB: 4 Gaps: 0

09-833799-13B (1-57) x PIGALP (1-600)

QY 5 VallysGlyProValSerThrLysProGlySerCysProIleLeuIleArgCysAla 24
Db 224 ATCAGGAACCCAGTTAAGCTGAAGCTGGGAAGTGTCCAGTGGTCTATGGCCAGTGTATG 283
QY 25 MetLeuAsnProProAsnArgCysLeuLysAspThrAspCysProGlyIleLysLysCys 44
Db 284 ATGCTCAACCCCCCAATCACTGCAAGACAGACAGCCAGTGCCTGGGTGACTTAAATGC 343
QY 45 CysGluGlySerCysGlyMetAlaCysPheValPro 56
Db 344 TGCAAGAGCATGTGCGGGAAGTGTGCTCACCCT 379

RESULT 39
MMU88093
LOCUS Mus musculus secretory leukocyte protease inhibitor mRNA, complete cds. ROD 19-APR-1997
DEFINITION
ACCESSION U88093
VERSION U88093.1 GI:1945382
KEYWORDS
SOURCE Mus musculus.
ORGANISM Mus musculus
REFERENCE Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
AUTHORS Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
TITLE 1 (bases 1 to 671)
Zitnik,R.J., Zhang,J., Kashem,M.A., Kohno,T., Lyons,D.E.,
Wright,C.D., Rosen,E., Goldberg,I. and Hayday,A.C.
The cloning and characterization of a murine secretory leukocyte
protease inhibitor cDNA
Biochem. Biophys. Res. Commun. 232 (3), 687-697 (1997)
JOURNAL
MEDLINE 97271386
PUBMED 9126337
REFERENCE 2 (bases 1 to 671)
AUTHORS Zitnik,R.J., Zhang,J., Kashem,M.A., Kohno,T., Lyons,D.E.,
Wright,C.D., Rosen,E., Goldberg,I. and Hayday,A.C.
Direct Submission
TITLE Submitted (31-JAN-1997) Internal Medicine, Yale University, 333
JOURNAL Cedar Street, New Haven, CT 06520, USA
FEATURES
source Location/Qualifiers
1..671
/organism="Mus musculus"
/db_xref="taxon:10090"
/cell_line="ras-transformed 3T3 cells"
13..408
/function="serine protease inhibitor"
/codon_start=1
/product="secretory leukocyte protease inhibitor"
/protein_id="AAC53140.1"
/db_xref="GI:1945383"
/translation="MKSCGLLPFTVLLALGILAPWTVEGGKNDKIGACPAKKPAQC
LKLEKPCQRTDWECPGKQCCODACGSKCVNPVPIRKPVRKRCVKTKQARCMMLNP
PNVCQRDQGDGKYKCEGICGKVCCLPPM"

BASE COUNT 157 a 179 c 177 g 158 t
ORIGIN
Alignment Scores:
Pred. No.: 2.24e-09 Length: 671
Score: 161.00 Matches: 28
Percent Similarity: 58.49% Conservative: 3
Best Local Similarity: 52.83% Mismatches: 22
Query Match: 49.24% Indels: 0
DB: 10 Gaps: 0

09-833799-13B (1-57) x MMU88093 (1-671)

QY 4 ProValLysGlyProValSerThrLysProGlySerCysProIleLeuIleArgCys 23
Db 244 CCCATTGCGAAACCAAGTGTGGAGGAAGCCTGGGAGGTGCGTCAAAACTCAGGCAAGATGT 303
QY 24 AlaMetLeuAsnProProAsnArgCysLeuLysAspThrAspCysProGlyIleLysLys 43
Db 304 ATGATGCTTTAACCTCCCAATGTCTGCAGAGGACGGGCGGAGTGTGACGCAATACAAG 363

QY 44 CysCysGluGlySerCysGlyMetAlaCysPheValPro 56
|||||
Db 364 TGCTGTGAGGGTATATGTGGAAAGTCTGCTGCCCCCG 402
RESULT 40
MMU94341
LOCUS
DEFINITION Mus musculus secretory leukoprotease inhibitor mRNA, complete cds.
ACCESSION U94341
VERSION U94341.1 GI:1945450
KEYWORDS
SOURCE Mus musculus.
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 682)
REFERENCE Abe, T., Tomimaga, Y., Kikuchi, T., Watanabe, A., Satoh, K., Watanabe, Y.
AUTHORS and Nukiwa, T.
TITLE Bacterial pneumonia causes augmented expression of the secretory
leukoprotease inhibitor gene in the murine lung
JOURNAL Am. J. Respir. Crit. Care Med. 156 (4 Pt 1), 1235-1240 (1997)
MEDLINE 98011992
PUBMED 9351627
REFERENCE 2 (bases 1 to 682)
AUTHORS Abe, T., Tomimaga, Y., Kikuchi, T., Watanabe, A., Satoh, K., Watanabe, Y.
and Nukiwa, T.
TITLE Direct Submission
JOURNAL Submitted (18-MAR-1997) Department of Respiratory Oncology and
Molecular Medicine, Institute of Development, Aging and Cancer,
Tohoku University, 4-1 Seiryomachi, Aoba-ku, Sendai 980-77, Japan
FEATURES
source
1. .682
/organism="Mus musculus"
/strain="B6CBA [C57Black/6 x CBA]"
/db_xref="taxon:10090"
/tissue_type="lung"
14. .409
/codon_start=1
/product="secretory leukoprotease inhibitor"
/protein_id="AAC53394.1"
/db_xref="GI:1945451"
/translation="MKSCGLLPFTVLLALGILAPWTVEGGKNDAIKIGACPAKKPAQC
LKLEKPCQRTDWECPGKQRCQDAGSKCVNPVPIRKPVRKPGRCVKVTKQARCMMLNP
PNVCQRDGGQCDGKYKCEGICGKVCCLPPM"
BASE COUNT 163 a 179 c 181 g 159 t
ORIGIN
Alignment Scores:
Pred. No.: 2.27e-09 Length: 682
Score: 161.00 Matches: 28
Percent Similarity: 58.49% Conservative: 3
Best Local Similarity: 52.83% Mismatches: 22
Query Match: 49.24% Indels: 0
DB: 10 Gaps: 0
09-833799-13B (1-57) x MMU94341 (1-682)
QY 4 ProValLysGlyProValSerThrLysProGlySerCysProIleLeuIleArgCys 23
|||||
Db 245 CCCATTGCGAAACACAGTGTGGAGGAGCCTGGAGGTGCTCAAAACTCAGGCAAGATGT 304
QY 24 AlaMetLeuAsnProProAsnArgCysLeuLysAspThrAspCysProGlyIleLysLys 43
|||||
Db 305 ATGATGCTTAACCTCCCAATGTCTGCCAGAGGACGGGACGAGTGTGACGGCAATAACAAG 364
QY 44 CysCysGluGlySerCysGlyMetAlaCysPheValPro 56
|||||
Db 365 TGCTGTGAGGGTATATGTGGAAAGTCTGCCTGCCCCCG 403
RESULT 41
BC028509
LOCUS

DEFINITION Mus musculus, secretory leukocyte protease inhibitor, clone
MGC:41142 IMAGE:1513866, mRNA, complete cds.
ACCESSION BC028509
VERSION BC028509.1 GI:20306995
KEYWORDS MGC.
SOURCE house mouse.
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 894)
REFERENCE Strausberg, R.
AUTHORS Direct Submission
TITLE
JOURNAL Submitted (23-APR-2002) National Institutes of Health, Mammalian
Gene Collection (MGC), Cancer Genomics Office, National Cancer
Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
USA
REMARK NIH-MGC Project URL: http://mgc.nci.nih.gov
COMMENT Contact: MGC help desk
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: Marcello Bento Soares, Ph.D.
cDNA Library Preparation: Soares Laboratory
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Genome Sequence Centre,
BC Cancer Agency, Vancouver, BC, Canada
info@bcgsc.bc.ca
Steven Jones, Jennifer Asano, Ian Bosdet, Yaron Butterfield,
Susanna Chan, Readman Chiu, Chris Fjell, Erin Garland, Ran Guin,
Leticia Hsiao, Martin Krzywinski, Reta Kutsche, Oliver Lee, Soo
Sen Lee, Victor Ling, Carrie Mathewson, Candice McLeavy, Steven
Ness, Pawan Pandoh, Anna-Lilisa Prabhu, Parvaneh Saeedi, Jacqueline
Schein, Duane Smalus, Michael Smith, Lorraine Spence, Jeff Stott,
Michael Thorne, Miranada Tsai, Natasja van den Bosch, Jill Vardy,
George Yang, Scott Zuyderduyn, Marco Marra.
Clone distribution: MGC clone distribution information can be found
through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov
Series: IRAK Plate: 67 Row: j Column: 6
This clone was selected for full length sequencing because it
passed the following selection criteria: matched mRNA gi: 6755573.
FEATURES
source
1. .894
/organism="Mus musculus"
/db_xref="taxon:10090"
/clone="MGC:41142 IMAGE:1513866"
/tissue_type="Mammary gland, lactating mouse"
/clone_lib="Soares_mammary_gland_NMLMG"
/lab_host="DH10B"
/note="Vector: pT7T3D-Pac"
236. .631
/codon_start=1
/product="secretory leukocyte protease inhibitor"
/protein_id="AAH28509.1"
/db_xref="GI:20306996"
/db_xref="LocusID:20568"
/translation="MKSCGLLPFTVLLALGILAPWTVEGGKNDAIKIGACPAKKPAQC
LKLEKPCQRTDWECPGKQRCQDAGSKCVNPVPIRKPVRKPGRCVKVTKQARCMMLNP
PNVCQRDGGQCDGKYKCEGICGKVCCLPPM"
BASE COUNT 205 a 242 c 233 g 214 t
ORIGIN
Alignment Scores:
Pred. No.: 3.02e-09 Length: 894
Score: 161.00 Matches: 28
Percent Similarity: 58.49% Conservative: 3
Best Local Similarity: 52.83% Mismatches: 22
Query Match: 49.24% Indels: 0
DB: 10 Gaps: 0
09-833799-13B (1-57) x BC028509 (1-894)
QY 4 ProValLysGlyProValSerThrLysProGlySerCysProIleLeuIleArgCys 23
|||||
Db 467 CCCATTGCGAAACACAGTGTGGAGGAGCCTGGAGGTGCTCAAAACTCAGGCAAGATGT 526

```
QY 24 AlaMetLeuAsnProProAsnArgCysLeuLysAspThrAspCysProGlyIleLysLys 43
|||||
Db 527 ATGATGCTTAACCTCCCAATGCTGCGAGAGGACGGCAGTGTGACGGCAATACAAG 586
|||||

QY 44 CysCysGluGlySerCysGlyMetAlaCysPheValPro 56
|||||
Db 587 TGCTGTGAGGGTATATGTGGGAAAGTCTGCCTGCCCCCG 625
|||||

RESULT 42
MMU73004
LOCUS
DEFINITION Mus musculus secretory leukocyte protease inhibitor mRNA, complete
cgs.
ACCESSION U73004
VERSION U73004.1 GI:1763262
KEYWORDS
SOURCE Mus musculus.
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 1123)
AUTHORS Jin,F.Y., Nathan,C., Radzioch,D. and Ding,A.
TITLE Secretory leukocyte protease inhibitor: a macrophage product
induced by and antagonistic to bacterial lipopolysaccharide
JOURNAL Cell 88 (3), 417-426 (1997)
MEDLINE 97191310
PUBMED 9039268
REFERENCE 2 (bases 1 to 1123)
AUTHORS Ding,A., Jin,F.-Y. and Nathan,C.F.
TITLE Direct Submission
JOURNAL Submitted (01-OCT-1996) Medicine, Cornell University Medical
College, 1300 York Ave. Box 57, New York, NY 10021, USA
FEATURES
source
1..1123
/organism="Mus musculus"
/db_xref="taxon:10090"
/cell_line="RAW 264.7"
/cell_type="macrophage"
447..842
/function="serine protease inhibitor"
/codon_start=1
/product="secretory leukocyte protease inhibitor"
/protein_id="AAC53047.1"
/db_xref="GI:1763263"
/translation="MKSCGLLPFTVLLALGILAPWTVEGGKNDAIKIGACPAKKPAQC
LKLEKPCQRTDWECPGKQRCQDACGSKVNPVPIRKPVWRKPGRCVKTQARCMMLNP
PNVCQDGGQCDGKYKCEGICGKVCPLPM"
BASE COUNT 286 a 297 c 288 g 252 t
ORIGIN

Alignment Scores:
Pred. No.: 3.84e-09 Length: 1123
Score: 161.00 Matches: 28
Percent Similarity: 58.49% Conservative: 3
Best Local Similarity: 52.83% Mismatches: 22
Query Match: 49.24% Indels: 0
DB: 10 Gaps: 0

09-833799-13B (1-57) x MMU73004 (1-1123)

QY 4 ProValLysGlyProValSerThrLysProGlySerCysProIleIleLeuIleArgCys 23
|||||
Db 678 CCCATTGCAAAACAGTGTGGAGGAGCGCTGGAGGTGCGTCAAAACTCAGGCAAGATGT 737
|||||

QY 24 AlaMetLeuAsnProProAsnArgCysLeuLysAspThrAspCysProGlyIleLysLys 43
|||||
Db 738 ATGATGCTTAACCTCCCAATGCTGCGAGAGGACGGCAGTGTGACGGCAATACAAG 797
|||||

QY 44 CysCysGluGlySerCysGlyMetAlaCysPheValPro 56
|||||
Db 798 TGCTGTGAGGGTATATGTGGGAAAGTCTGCCTGCCCCCG 836
|||||
```

```
RESULT 43
AF151982
LOCUS
DEFINITION Rattus norvegicus secretory leukocyte protease inhibitor (SLPI)
mRNA, complete cds.
ACCESSION AF151982
VERSION AF151982.1 GI:4929548
KEYWORDS
SOURCE Rattus norvegicus.
ORGANISM Rattus norvegicus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
REFERENCE 1 (bases 1 to 667)
AUTHORS Chen,D.H., Xu,X.P., Bagchi,M.K. and Bagchi,I.C.
TITLE Molecular cloning and spatio-temporal expression of rat secretory
leukocyte protease inhibitor (SLPI) in the uterus
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 667)
AUTHORS Chen,D.H., Xu,X.P., Bagchi,M.K. and Bagchi,I.C.
TITLE Direct Submission
JOURNAL Submitted (17-MAY-1999) Center for Biomedical Research, The
Population Council, 1230 York Avenue, New York, NY 10021, USA
FEATURES
source
1..667
/organism="Rattus norvegicus"
/strain="SD"
/db_xref="taxon:10116"
1..667
/gene="SLPI"
13..405
/gene="SLPI"
/function="serine protease inhibitor"
/codon_start=1
/product="secretory leukocyte protease inhibitor"
/protein_id="AAD34035.1"
/db_xref="GI:4929549"
/translation="MKSSGLFPLMVLALGVLPWSVEGGKNDAIKIGACPAKKPAQC
LKREKPECSTDWCPGKQRCQDTCGFKCLNPVPIRGPVKPGRCCLKFQKCLMLNPP
NKCQNDGQCDGKYKCEGMCCKVCLPPV"
BASE COUNT 176 a 159 c 161 g 171 t
ORIGIN

Alignment Scores:
Pred. No.: 2.53e-09 Length: 667
Score: 160.50 Matches: 28
Percent Similarity: 60.38% Conservative: 4
Best Local Similarity: 52.83% Mismatches: 20
Query Match: 49.08% Indels: 1
DB: 10 Gaps: 1

09-833799-13B (1-57) x AF151982 (1-667)

QY 4 ProValLysGlyProValSerThrLysProGlySerCysProIleIleLeuIleArgCys 23
|||||
Db 244 CCCATTGCTGGACCAAGTGAAG---AAGCCTGGGAGGTGCTCAAAATTTCAAGGAAATGT 300
|||||

QY 24 AlaMetLeuAsnProProAsnArgCysLeuLysAspThrAspCysProGlyIleLysLys 43
|||||
Db 301 CTGATGCTTAACCTCCCAATAAGTCCAGAAATGACGGCCAGTGTGATGGCAATACAAA 360
|||||

QY 44 CysCysGluGlySerCysGlyMetAlaCysPheValPro 56
|||||
Db 361 TGTGTGAGGGTATGTGTGGGAAAGTCTGCCTTCCCCCA 399
|||||

RESULT 44
PIGWAPC
LOCUS
DEFINITION Pig DNA for elafin family member protein, complete cds.
ACCESSION D50321
VERSION D50321.1 GI:1228061
KEYWORDS elafin family member protein.
SOURCE Sus scrofa DNA, clone_lib:EMBL SP6/T7 clone:lambda WAP-3.
```

ORGANISM Sus scrofa
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
REFERENCE 1 (sites)
AUTHORS Tamechika, I., Itakura, M., Saruta, Y., Furukawa, M., Kato, A.,
Tachibana, S. and Hirose, S.
TITLE Accelerated evolution in inhibitor domains of porcine elafin family
members
JOURNAL J. Biol. Chem. 271 (12), 7012-7018 (1996)
MEDLINE 96215132
REFERENCE 2 (bases 1 to 3670)
AUTHORS Hirose, S.
JOURNAL Unpublished
REFERENCE 3 (bases 1 to 3670)
AUTHORS Hirose, S.
TITLE Direct Submission
JOURNAL Submitted (18-APR-1995) Shigehisa Hirose, Tokyo Institute of
Technology, Department of Biological Sciences; 4259 Nagatsuta-cho,
Midori-ku, Yokohama, Kanagawa 226-8501, Japan
(E-mail: shirose@bio.titech.ac.jp, Tel: 045-924-5726,
Fax: 045-924-5824)

FEATURES
source Location/Qualifiers
1..3670
/organism="Sus scrofa"
/db_xref="taxon:9823"
/clone="lambda WAP-3"
/clone_lib="EMBL SP6/T7"
<1..710
/number=1
<1..628
join(629..710,1552..1904)
/codon_start=1
/product="elafin family member protein"
/protein_id="BAA08856.1"
/db_xref="GI:1228062"
translation="MRSRSFLVLVAVFLICETLVAQRLDIRGPKGQGQDPVEGQDQD
EGQGPVKVEILDIGQDLVKGQDPVEGQDPVKAQLPDKVQDPVKAQPIIQGFLFPKPG
VCPKIIIFCPLVNPPIKCWRDSHCPGVKKCCPSLCGKGCVTPR"
629..691
1722..1901
/product="unnamed"
711..1551
/number=1
1552..1905
/number=2
join(1905,2414..2570)
1906..2413
/number=2
repeat_unit 2110..2371
/note="SINE (PRE-1)"
exon 2414..2570
/number=3
polyA_signal 2549..2554
polyA_site 2570
repeat_region 3581..3630
/note="TG repeat"
BASE COUNT 774 a 951 c 1131 g 814 t
ORIGIN

Alignment Scores:
Pred. No.: 1.73e-08 Length: 3670
Score: 160.00 Matches: 27
Percent Similarity: 65.52% Conservative: 11
Best Local Similarity: 46.55% Mismatches: 18
Query Match: 48.93% Indels: 2
DB: 4 Gaps: 2
09-833799-13B (1-57) x PIGWAPC (1-3670)
Qy 1 AlaGlnGluProValLysGlyPro---ValSerThrLysProGlySerCysProIleIle 19
Db 1731 GCGCAACCTCCAATCCAAGGAGGATTCCTTTTCCCTAAGCCAGGCGTCTGCCCCAAGATT 1790

Qy 20 LeulleArgCysAlaMetLeuAsnProProAsnArgCysLeuLysAspThrAspCysPro 39
Db 1791 ATCTTT---TGCCCGCTGGTCAATCCCCCTATCAAGTGTGGAGAGATTCTCACTGCCCA 1847
Qy 40 GlyTleLysLysCysCysGluGlySerCysGlyMetAlaCysPheValProGln 57
Db 1848 GGGGTCAAGAAAGTGTCTGCCAAGCCTTTGCGGGAAGGATGTGTGACTCCCAGG 1901
RESULT 45
AB042257
LOCUS AB042257 464 bp mRNA linear ROD 30-OCT-2001
DEFINITION Cavia porcellus mRNA for caltrin-like protein II, complete cds.
ACCESSION AB042257
VERSION AB042257.1 GI:16519354
KEYWORDS
SOURCE Cavia porcellus tissue_lib:seminal vesicle cDNA to mRNA.
ORGANISM Cavia porcellus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Hystricognathi; Caviidae; Cavia.
REFERENCE 1
AUTHORS Coronel, C.E., San Agustín, J. and Lardy, H.A.
TITLE Purification and structure of caltrin-like proteins from seminal
vesicle of the guinea pig
JOURNAL J. Biol. Chem. 265 (12), 6854-6859 (1990)
MEDLINE 90216715
REFERENCE 2
AUTHORS Furutani, Y. and Hirose, S.
TITLE Evolution of caltrin-like protein
JOURNAL Unpublished
REFERENCE 3 (bases 1 to 464)
AUTHORS Furutani, Y. and Hirose, S.
TITLE Direct Submission
JOURNAL Submitted (26-APR-2000) Shigehisa Hirose, Tokyo Institute of
Technology, Department of Bioscience; Nagatadacho4259 Midori-ku,
Yokohama, Kanagawa 226-8501, Japan
(E-mail: shirose@bio.titech.ac.jp, Tel: 81-45-924-5726,
Fax: 81-45-924-5824)

FEATURES
source Location/Qualifiers
1..464
/organism="Cavia porcellus"
/db_xref="taxon:10141"
/tissue_lib="seminal vesicle"
17..247
/note="antileukoproteinase repeat homology #label ALP"
/codon_start=1
/product="caltrin-like protein II"
/protein_id="BAB70710.1"
/db_xref="GI:16519355"
translation="MKATILFILFLILEKPSFGRRLLHGQAINRPGSCPVMYICPP
RHPPNKCTSDYDCPKQKCCPGYCGKQCYQE"
17..79
sig_peptide
mat_peptide 80..244
/product="unnamed"
/db_xref="SWISS-PROT:P22075"
misc_feature 110..247
/note="WAP. domain"
BASE COUNT 113 a 139 c 97 g 115 t
ORIGIN

Alignment Scores:
Pred. No.: 3.33e-09 Length: 464
Score: 158.00 Matches: 24
Percent Similarity: 62.75% Conservative: 8
Best Local Similarity: 47.06% Mismatches: 19
Query Match: 48.32% Indels: 0
DB: 10 Gaps: 0
09-833799-13B (1-57) x AB042257 (1-464)
Qy 7 GlyProValSerThrLysProGlySerCysProIleIleLeuIleArgCysAlaMetLeu 26
Db 92 GGGCAGCCCATCAATAGGCTTGGTTCCTGTCCCCGGGTATGATATACTGCCCCGCAAGG 151

Query Match: 44.04% Indels: 0
DB: 10 Gaps: 0

09-833799-13B (1-57) x AF205374 (1-4090)

Qy 12 LysProGlySerCysProIleIleLeuIleArgCysAlaMetLeuAsnProProAsnArg 31
|||||
Db 2570 AAGCCTGGGAGGTGCGTCAAAACTCAGGCAAGATGATGATGCTTAACCTCCCAATGTC 2629

Qy 32 CysLeuLysAspThrAspCysProGlyIleLysLysCysCysGluGlySerCysGlyMet 51
|||
Db 2630 TGCCAGAGGACGGGCGAGTGTGACGGCAAAATACAAGTGTGTGAGGGTATATGTGGGAAA 2689

Qy 52 AlaCysPheValPro 56
|||
Db 2690 GTCTGCCTGCCCCCG 2704

RESULT 50
AR067991

LOCUS AR067991 194 bp DNA linear PAT 29-SEP-1999
DEFINITION Sequence 5 from patent US 5851983.
ACCESSION AR067991
VERSION AR067991.1 GI:5999213
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
UNCLASSIFIED.

REFERENCE 1 (bases 1 to 194)
AUTHORS Sugiyama,T., Kamimura,T., Masuda,K., Okada,M., Ohtsuka,E.,
Imaizumi,A., Watanabe,K., Suga,T., Matsumoto,Y. and Takeuchi,A.
TITLE Elastase inhibitory polypeptide and process for production thereof
by recombinant gene technology
JOURNAL Patent: US 5851983-A 5 22-DEC-1998;
FEATURES Location/Qualifiers
source 1..194
BASE COUNT 48 a 42 c 53 g 51 t
ORIGIN

Alignment Scores:
Pred. No.: 6.76e-08 Length: 194
Score: 143.00 Matches: 27
Percent Similarity: 55.36% Conservative: 4
Best Local Similarity: 48.21% Mismatches: 23
Query Match: 43.73% Indels: 2
DB: 6 Gaps: 1

09-833799-13B (1-57) x AR067991 (1-194)

Qy 3 GluProValLysGlyProValSerThr-----LysProGlySerCysProIleIleLeu 20
:::|||||
Db 1 GATCCGGTCGACACCCCGAACCCGACGCGTCGTAAACCGGGTAAATGTCGGTTACATAT 60

Qy 21 IleArgCysAlaMetLeuAsnProProAsnArgCysLeuLysAspThrAspCysProGly 40
:::|||||
Db 61 GGTCAAGTGTCTGATGCTGAACCCCGCCGAACCTTCTGTGAAATGGACGGTCAAGTGTAAACGA 120

Qy 41 IleLysLysCysCysGluGlySerCysGlyMetAlaCysPheValPro 56
|||||
Db 121 GATCTGAAATGTTGATGGGTATGTGTGTTAAATCTGTGTTTCTCCG 168

RESULT 51
I08404

LOCUS I08404 324 bp DNA linear PAT 02-DEC-1994
DEFINITION Sequence 1 from Patent WO 8603519.
ACCESSION I08404
VERSION I08404.1 GI:588883
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
UNCLASSIFIED.

REFERENCE 1 (bases 1 to 324)
AUTHORS Bandyopadhyay,P.K., Eisenberg,S.P., Stetler,G.L. and Thompson,R.C.

TITLE RECOMBINANT METHODS FOR PRODUCTION OF SERINE PROTEASE INHIBITORS
AND DNA SEQUENCES USEFUL FOR SAME
JOURNAL Patent: WO 8603519-A 1 19-JUN-1986;
FEATURES Location/Qualifiers
source 1..324
BASE COUNT 87 a 85 c 88 g 64 t
ORIGIN

Alignment Scores:
Pred. No.: 1.16e-07 Length: 324
Score: 143.00 Matches: 27
Percent Similarity: 55.36% Conservative: 4
Best Local Similarity: 48.21% Mismatches: 23
Query Match: 43.73% Indels: 2
DB: 6 Gaps: 1

09-833799-13B (1-57) x I08404 (1-324)

Qy 3 GluProValLysGlyProValSerThr-----LysProGlySerCysProIleIleLeu 20
:::|||||
Db 145 GATCCGGTTGATACCCCGAACCCGACTCGTCGAAACCGGTAATGCCCGGTAACTTAT 204

Qy 21 IleArgCysAlaMetLeuAsnProProAsnArgCysLeuLysAspThrAspCysProGly 40
:::|||||
Db 205 GGCCAGTGTCTGATGCTGAACCCCGCCGAACCTTCTGCGAAATGGACGGCCAGTGTAAACGA 264

Qy 41 IleLysLysCysCysGluGlySerCysGlyMetAlaCysPheValPro 56
|||||
Db 265 GATCTGAAATGCTGTATGGGTATGTGCGCAAAATCTGTGTTTCCCG 312

RESULT 52
AX302535

LOCUS AX302535 399 bp DNA linear PAT 30-NOV-2001
DEFINITION Sequence 53 from Patent WO0175177.
ACCESSION AX302535
VERSION AX302535.1 GI:17383077
KEYWORDS
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1
AUTHORS Morin,P.J., Sherman-Baust,C.A., Pizer,E.S. and Hough,C.D.
TITLE Tumor markers in ovarian cancer
JOURNAL Patent: WO 0175177-A 53 11-OCT-2001;
THE SECRETARY OF THE DEPARTMENT OF HEALTH AND HUMAN SERVICES (US)

FEATURES Location/Qualifiers
source 1..399
/organism="Homo sapiens"
/db_xref="taxon:9606"

BASE COUNT 91 a 100 c 111 g 97 t
ORIGIN

Alignment Scores:
Pred. No.: 1.44e-07 Length: 399
Score: 143.00 Matches: 27
Percent Similarity: 55.36% Conservative: 4
Best Local Similarity: 48.21% Mismatches: 23
Query Match: 43.73% Indels: 2
DB: 6 Gaps: 1

09-833799-13B (1-57) x AX302535 (1-399)

Qy 3 GluProValLysGlyProValSerThr-----LysProGlySerCysProIleIleLeu 20
:::|||||
Db 220 GATCCTGTTGACACCCCAACCAACAGGAGGAGCCTGGGAAGTGGCCAGTGACTTAT 279

Qy 21 IleArgCysAlaMetLeuAsnProProAsnArgCysLeuLysAspThrAspCysProGly 40
:::|||||
Db 280 GGCCAATGTTTGTATGCTTAACCCCCCAATTTCTGTGAGATGGATGGCCAGTGCAGCGT 339

Qy 41 IleLysLysCysCysGluGlySerCysGlyMetAlaCysPheValPro 56

NFCEMDGQCKRDLKCCMGMCCKSCVSPVKA"

```
BASE COUNT      106 a      162 c      171 g      126 t
ORIGIN

Alignment Scores:
Pred. No.:      2.07e-07      Length:      565
Score:          143.00      Matches:      27
Percent Similarity: 55.36%      Conservative: 4
Best Local Similarity: 48.21%      Mismatches: 23
Query Match:    43.73%      Indels: 2
DB:              6      Gaps: 1

09-833799-13B (1-57) x AX069252 (1-565)

QY      3 GluProValLysGlyProValSerThr-----LysProGlySerCysProIleLeu 20
      ::::::::::|
Db      278 GATCCTGTTGACACCCCAACCAAGGAGGAGCTGGGAAGTGGCCAGTGACTTAT 337
      GATCCTGTTGACACCCCAACCAAGGAGGAGCTGGGAAGTGGCCAGTGACTTAT 337

QY      21 IleArgCysAlaMetLeuAsnProProAsnArgCysLeuLysAspThrAspCysProGly 40
      ::::::::::|
Db      338 GGCCAATGTTTGATGCTTAACCCCCCAATTTCTGTGAGATGGATGGCCAGTGCAAGCGT 397
      GGCCAATGTTTGATGCTTAACCCCCCAATTTCTGTGAGATGGATGGCCAGTGCAAGCGT 397

QY      41 IleLysLysCysCysGluGlySerCysGlyMetAlaCysPheValPro 56
      ::::::::::|
Db      398 GACTTGAAGTGTTCATGGGCATGTGTGGAAATCTCGCTTTCCCT 445
      GACTTGAAGTGTTCATGGGCATGTGTGGAAATCTCGCTTTCCCT 445

RESULT 56
I01502
LOCUS
DEFINITION      Sequence 3 from Patent US 4845076.
ACCESSION      I01502
VERSION      I01502.1 GI:270158
KEYWORDS
SOURCE
ORGANISM
REFERENCE      1 (bases 1 to 565)
AUTHORS      Heinzl, R., Appelhans, H., Gassen, H.G., Machleidt, W. and Seemuller, U.
TITLE      DNA sequences coding for proteins having the biological activity of HUSI-type I inhibitors, biotechnological methods for the preparation of said proteins and pharmaceutical compositions containing said proteins
JOURNAL      Patent: US 4845076-A 3 04-JUL-1989; Grunenthal GmbH;;
DE;

FEATURES
source
Location/Qualifiers
BASE COUNT      106 a      162 c      171 g      126 t
ORIGIN

Alignment Scores:
Pred. No.:      2.07e-07      Length:      565
Score:          143.00      Matches:      27
Percent Similarity: 55.36%      Conservative: 4
Best Local Similarity: 48.21%      Mismatches: 23
Query Match:    43.73%      Indels: 2
DB:              6      Gaps: 1

09-833799-13B (1-57) x I01502 (1-565)

QY      3 GluProValLysGlyProValSerThr-----LysProGlySerCysProIleLeu 20
      ::::::::::|
Db      278 GATCCTGTTGACACCCCAACCAAGGAGGAGCTGGGAAGTGGCCAGTGACTTAT 337
      GATCCTGTTGACACCCCAACCAAGGAGGAGCTGGGAAGTGGCCAGTGACTTAT 337

QY      21 IleArgCysAlaMetLeuAsnProProAsnArgCysLeuLysAspThrAspCysProGly 40
      ::::::::::|
Db      338 GGCCAATGTTTGATGCTTAACCCCCCAATTTCTGTGAGATGGATGGCCAGTGCAAGCGT 397
      GGCCAATGTTTGATGCTTAACCCCCCAATTTCTGTGAGATGGATGGCCAGTGCAAGCGT 397

QY      41 IleLysLysCysCysGluGlySerCysGlyMetAlaCysPheValPro 56
      ::::::::::|
Db      398 GACTTGAAGTGTTCATGGGCATGTGTGGAAATCTCGCTTTCCCT 445
      GACTTGAAGTGTTCATGGGCATGTGTGGAAATCTCGCTTTCCCT 445
```

```
RESULT 57
HSSLIPR
LOCUS
DEFINITION      Human SLPI mRNA fragment for secretory leucocyte protease inhibitor.
ACCESSION      X04503
VERSION      X04503.1 GI:36490
KEYWORDS      elastase inhibitor; protease inhibitor; secretory leucocyte protease inhibitor; trypsin inhibitor.
SOURCE      Homo sapiens.
ORGANISM      Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE      1 (bases 1 to 573)
AUTHORS      Stetler, G., Brewer, M.T. and Thompson, R.C.
TITLE      Isolation and sequence of a human gene encoding a potent inhibitor of leukocyte proteases
JOURNAL      Nucleic Acids Res. 14 (20), 7883-7896 (1986)
MEDLINE      87040761
PUBMED      3640338
COMMENT      Data kindly reviewed (15-SEP-1987) by Stetler G.
FEATURES
source
Location/Qualifiers
1..573
/organism="Homo sapiens"
/db_xref="taxon:9606"
/tissue_type="parotid gland"
1..399
/note="SLPI-precursor"
/codon_start=1
/protein_id="CAA28188.1"
/db_xref="GI:36491"
/db_xref="SWISS-PROT:P03973"
/translation="MKSSGLFPFLVLLALGTLPWAVEGSGKSFKAGVCPPKKSAOCLRYKKPECCSDWQCPGKKRCCPDTCGIKCLDPDVTNPTRRKPGKCPVTYGGQCLMLNPPNFCEMDGQCKRDLKCCMGMCCKSCVSPVKA"
1..75
sig_peptide
mat_peptide
76..396
/product="mature SLPI (AA 1-107)"
/misc_feature
546..552
/note="put. polyA signal"
573
polyA_site
/note="polyA site"
BASE COUNT      131 a      145 c      152 g      145 t
ORIGIN

Alignment Scores:
Pred. No.:      2.11e-07      Length:      573
Score:          143.00      Matches:      27
Percent Similarity: 55.36%      Conservative: 4
Best Local Similarity: 48.21%      Mismatches: 23
Query Match:    43.73%      Indels: 2
DB:              9      Gaps: 1

09-833799-13B (1-57) x HSSLIPR (1-573)

QY      3 GluProValLysGlyProValSerThr-----LysProGlySerCysProIleLeu 20
      ::::::::::|
Db      220 GATCCTGTTGACACCCCAACCAAGGAGGAGCTGGGAAGTGGCCAGTGACTTAT 279
      GATCCTGTTGACACCCCAACCAAGGAGGAGCTGGGAAGTGGCCAGTGACTTAT 279

QY      21 IleArgCysAlaMetLeuAsnProProAsnArgCysLeuLysAspThrAspCysProGly 40
      ::::::::::|
Db      280 GGCCAATGTTTGATGCTTAACCCCCCAATTTCTGTGAGATGGATGGCCAGTGCAAGCGT 339
      GGCCAATGTTTGATGCTTAACCCCCCAATTTCTGTGAGATGGATGGCCAGTGCAAGCGT 339

QY      41 IleLysLysCysCysGluGlySerCysGlyMetAlaCysPheValPro 56
      ::::::::::|
Db      340 GACTTGAAGTGTTCATGGGCATGTGTGGAAATCTCGCTTTCCCT 387
      GACTTGAAGTGTTCATGGGCATGTGTGGAAATCTCGCTTTCCCT 387

RESULT 58
AX328399
LOCUS
DEFINITION      Sequence 1 from Patent WO0190421.
ACCESSION      AX328399
AX328399
Sequence 1 from Patent WO0190421.
linear
DNA
594 bp
PAT 07-JAN-2002
```

09-833799-13B (1-57) x AX334507 (1-594)

QY 3 GluProValLysGlyProValSerThr-----LysProGlySerCysProIleLeu 20
 :::||||| ||| ||||| ||||| ||||| ||||| :::

Db 238 GATCCTGTTGACACCCCAAAACCAAGGAGGAAGCCTGGGAAGTGCCCAAGTACTTAT 297

QY 21 IleArgCysAlaMetLeuAsnProProAsnArgCysLeuLysAspThrAspCysProGly 40
 :::||||| ||||| ||||| ||||| ||||| ||||| |||||

Db 298 GGCCAATGTTTGATGTCTTAACCCCCCAATTCTGTGAGATGGATGGCCAGTGCAAGCGT 357

QY 41 IleLysLysCysCysGluGlySerCysGlyMetAlaCysPheValPro 56
 ||||| ||||| ||||| ||||| ::||| ||||| |||||

Db 358 GACTTGAAGTGTGCATGGCATGTGTGGAAAATCCTGCGTTTCCCCT 405

RESULT 60
AX335376

LOCUS AX335376 594 bp DNA linear PAT 09-JAN-2002

DEFINITION Sequence 5885 from Patent WO0194629.

ACCESSION AX335376

VERSION AX335376.1 GI:18126095

KEYWORDS

SOURCE human.

ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE

AUTHORS Young,P.E., Augustus,M., Carter,K.C., Ebner,R., Endress,G.,
Horrigan,S., Soppet,D.R. and Weaver,Z.

TITLE Cancer gene determination and therapeutic screening using signature
gene sets

JOURNAL Patent: WO 0194629-A 5885 13-DEC-2001;

FEATURES

source Avalon Pharmaceuticals (US)
 Location/Qualifiers
 1..594
 /organism="Homo sapiens"
 /db_xref="taxon:9606"

BASE COUNT 132 a 156 c 155 g 151 t

ORIGIN

Alignment Scores:
Pred. No.: 2.19e-07 Length: 594
Score: 143.00 Matches: 27
Percent Similarity: 55.36% Conservative: 4
Best Local Similarity: 48.21% Mismatches: 23
Query Match: 43.73% Indels: 2
DB: 6 Gaps: 1

09-833799-13B (1-57) x AX335376 (1-594)

QY 3 GluProValLysGlyProValSerThr-----LysProGlySerCysProIleLeu 20
 :::||||| ||| ||||| ||||| ||||| ||||| :::

Db 238 GATCCTGTTGACACCCCAAAACCAAGGAGGAAGCCTGGGAAGTGCCCAAGTACTTAT 297

QY 21 IleArgCysAlaMetLeuAsnProProAsnArgCysLeuLysAspThrAspCysProGly 40
 :::||||| ||||| ||||| ||||| ||||| ||||| |||||

Db 298 GGCCAATGTTTGATGTCTTAACCCCCCAATTCTGTGAGATGGATGGCCAGTGCAAGCGT 357

QY 41 IleLysLysCysCysGluGlySerCysGlyMetAlaCysPheValPro 56
 ||||| ||||| ||||| ||||| ::||| ||||| |||||

Db 358 GACTTGAAGTGTGCATGGCATGTGTGGAAAATCCTGCGTTTCCCCT 405

RESULT 61
HSALPR

LOCUS HSALPR 594 bp mRNA linear PRI 21-MAR-1995

DEFINITION Human mRNA for antileukoprotease (ALP) from cervix uterus.

ACCESSION X04470

VERSION X04470.1 GI:28638

KEYWORDS antileukoprotease; elastase inhibitor; protease; signal peptide.

SOURCE Homo sapiens.

ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

AUTHORS Heinzl,R., Appelhans,H., Gassen,G., Seemuller,U., Machleidt,W.,
Fritz,H. and Steffens,G.
TITLE Molecular cloning and expression of cDNA for human
antileukoprotease from cervix uterus
JOURNAL Eur. J. Biochem. 160 (1), 61-67 (1986)
MEDLINE 87030258
PUBMED 3533531
COMMENT Data kindly reviewed (05-DEC-1986) by H. Appelhans.
FEATURES
source Location/Qualifiers
1..594
/organism="Homo sapiens"
/db_xref="taxon:9606"
19..417
/note="precursor ALP"
/codon_start=1
/protein_id="CAA28158.1"
/db_xref="GI:28639"
/db_xref="SWISS-PROT:P03973"
/translation="MKSSGLFPFLVLLALGTLAPWAVEGSGKSFKAGVCPPPKKSAQCL
RYKPEQSDWQCPGKRCPCPDTCGKICLDPVDTNPTRRKPGKCPVTYGGQCLMLNPP
NFCMDGQCKRDLKCCMGCMGKSCVSPVKA"
19..93
/note="putative"
94..414
/product="put. mature peptide (aa 1-107)"
misc_feature 564..569
/note="pot. polyA signal"
misc_feature 568..573
/note="pot. polyA signal"
BASE COUNT 132 a 156 c 155 g 151 t
ORIGIN
Alignment Scores:
Pred. No.: 2.19e-07 Length: 594
Score: 143.00 Matches: 27
Percent Similarity: 55.36% Conservative: 4
Best Local Similarity: 48.21% Mismatches: 23
Query Match: 43.73% Indels: 2
DB: 9 Gaps: 1
09-833799-13B (1-57) x HSLPR (1-594)
QY 3 GluProValLysGlyProValSerThr-----LysProGlySerCysProIleLeu 20
Db 238 GATCCTGTTGACACCCCAACCAAGGAGGAGCCTGGGAAGTGCCAGTGACTTAT 297
QY 21 IleArgCysAlaMetLeuAsnProProAsnArgCysLeuLysAspThrAspCysProGly 40
Db 298 GGCCAATGTTGATGCTTAACCCCAATTTCTGTGAGATGGATGGCAGTCAAGCGT 357
QY 41 IleLysLysCysCysGluGlySerCysGlyMetAlaCysPheValPro 56
Db 358 GACTTGAAGTGTTCATGGGCATGTGTGGGAAATCCTGCGTTTCCCT 405
RESULT 62
AX014898
LOCUS AX014898 599 bp DNA linear PAT 07-SEP-2000
DEFINITION Sequence 97 from Patent WO9953040.
ACCESSION AX014898
VERSION AX014898.1 GI:10041165
KEYWORDS
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 599)
AUTHORS Schmitt,A., Specht,T., Dahl,E., Hinzmann,B., Rosenthal,A. and
Pilarczyk,C.
TITLE Human nucleic acid sequences from ovarian tumour tissue
JOURNAL Patent: WO 9953040-A 97 21-OCT-1999;
SCHMITT ARMIN (DE); SPECHT THOMAS (DE); DAHL EDGAR (DE); HINZMANN
BERND (DE); ROSENTHAL ANDRE (DE); METAGEN GES FUER GENOMFORSCHUN
(DE); PILARSKY CHRISTIAN (DE)

FEATURES
source Location/Qualifiers
1..599
/organism="Homo sapiens"
/db_xref="taxon:9606"
142 a 153 c 154 g 150 t
ORIGIN
Alignment Scores:
Pred. No.: 2.21e-07 Length: 599
Score: 143.00 Matches: 27
Percent Similarity: 55.36% Conservative: 4
Best Local Similarity: 48.21% Mismatches: 23
Query Match: 43.73% Indels: 2
DB: 6 Gaps: 1
09-833799-13B (1-57) x AX014898 (1-599)
QY 3 GluProValLysGlyProValSerThr-----LysProGlySerCysProIleLeu 20
Db 233 GATCCTGTTGACACCCCAACCAAGGAGGAGCCTGGGAAGTGCCAGTGACTTAT 292
QY 21 IleArgCysAlaMetLeuAsnProProAsnArgCysLeuLysAspThrAspCysProGly 40
Db 293 GGCCAATGTTGATGCTTAACCCCAATTTCTGTGAGATGGATGGCAGTCAAGCGT 352
QY 41 IleLysLysCysCysGluGlySerCysGlyMetAlaCysPheValPro 56
Db 353 GACTTGAAGTGTTCATGGGCATGTGTGGGAAATCCTGCGTTTCCCT 400
RESULT 63
BC020708
LOCUS BC020708 625 bp mRNA linear PRI 22-JAN-2002
DEFINITION Homo sapiens, secretory leukocyte protease inhibitor
(antileukoproteinase), clone MGC:22479 IMAGE:4733996, mRNA,
complete cds.
ACCESSION BC020708
VERSION BC020708.1 GI:18088404
KEYWORDS MGC.
SOURCE
ORGANISM Homo sapiens.
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 625)
AUTHORS Strausberg,R.
TITLE Direct Submission
JOURNAL Submitted (03-JAN-2002) National Institutes of Health, Mammalian
Gene Collection (MGC), Cancer Genomics Office, National Cancer
Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
USA
REMARK NIH-MGC Project URL: <http://mgc.nci.nih.gov>
COMMENT Contact: MGC help desk
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: CLONTECH
cDNA Library Preparation: CLONTECH Laboratories, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Sequencing Group at the Stanford Human Genome
Center, Stanford University School of Medicine, Stanford, CA 94305
Web site: <http://www-shgc.stanford.edu>
Contact: (Dickson, Mark) mcd@paxil.stanford.edu
Dickson, M., Schmutz, J., Grimwood, J., Rodriguez, A., and Myers,
R. M.
Clone distribution: MGC clone distribution information can be found
through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>
Series: IRAL Plate: 37 Row: i Column: 13
This clone was selected for full length sequencing because it
passed the following selection criteria: matched mRNA gi: 15834622.
Location/Qualifiers
1..625
/organism="Homo sapiens"
/db_xref="LocusID:6590"
/db_xref="taxon:9606"
/clone="MGC:22479 IMAGE:4733996"


```
promoter      181. .184
              /note="put. CAAT-box"
promoter      242. .247
              /note="put. TATA-box"
precursor_RNA 276. .2574
              /note="put. primary transcript of SLPI"
mRNA          276. .375
              /note="put. exon 1"
              join(291. .375,1092. .1250,1668. .1817,2397. .2401)
              /codon_start=1
              /product="secretory leukocyte protease inhibitor (SLPI)"
              /protein_id="CAA28187.1"
              /db_xref="GI:758101"
              /db_xref="SWISS-PROT:P03973"
              /translation="MKSSGLFPFLVLLALGTLAPWAVEGSGSKSFKAGVCPPKKSAQCL
              RYKPEQSDWQCPGKRRCCPDTCGIKCLDPVDTNPTRRKPGKCPVTYGGCLMLNPP
              NFCEMDGQCKRDLKCCMGCMGKSCVSPVKA"
              291. .365
              /note="signal peptide (AA -25 to -1)"
              join(366. .375,1092. .1250,1668. .1817,2397. .2398)
              /gene="SLPI"
              /product="unnamed"
              join(366. .375,1092. .1250,1668. .1817,2397. .2398)
              /gene="SLPI"
              376. .1091
              /gene="SLPI"
              /note="intron I"
              1092. .1250
              /gene="SLPI"
              /note="exon 2"
              1251. .1667
              /gene="SLPI"
              /note="intron II"
              1667. .1817
              /gene="SLPI"
              /note="exon 3"
              1818. .2396
              /gene="SLPI"
              /note="intron III"
              2397. .2574
              /note="exon 4"
              2549. .2554
              /note="polyA signal"
              2574
              /note="polyA site"
BASE COUNT   545 a   649 c   704 g   759 t
ORIGIN
Alignment Scores:
Pred. No.:      1.78e-06      Length:      2657
Score:          141.00      Matches:      24
Percent Similarity: 55.10%      Conservative: 3
Best Local Similarity: 48.98%      Mismatches: 22
Query Match:     43.12%      Indels: 0
DB:              9      Gaps: 0

09-833799-13B (1-57) x HSSLIPG (1-2657)
QY      8 ProValSerThrLysProGlySerCysProIleIleLeuIleArgCysAlaMetLeuAsn 27
      ||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db      1664 CCAGCAAGGAGGAGCCTGGGAAGTGCCAGTGACCTATGGCCAATGTTTGATGCTTAAC 1723
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY      28 ProProAsnArgCysLeuLysAspThrAspCysProGlyIleLysLysCysCysGluGly 47
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db      1724 CCCCCCAATTCTGTGAGATGGATGGCCAGTGCAAGCGTGACTTGAAGTGTTCATGGGC 1783
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY      48 SerCysGlyMetAlaCysPheValPro 56
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db      1784 ATGTGTGGGAAATCCTGCGTTTCCCT 1810
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
RESULT 66
A08111
LOCUS      A08111      321 bp      DNA      linear      PAT 28-JUL-1993
```

```
DEFINITION   Synthetic ALP-gene 242.
ACCESSION    A08111
VERSION      A08111.1  GI:413358
KEYWORDS
SOURCE       synthetic construct.
ORGANISM     synthetic construct
              artificial sequences.
REFERENCE    1 (bases 1 to 321)
AUTHORS      Heinzel-Wieland,R., Ammann,J., Steffens,G.J. and Flohe,L.
TITLE        Serine protease inhibitor proteins, medicaments containing them,
              DNA sequences coding for these proteins and methods for producing
              these proteins, medicaments and DNA sequences
              Patent: EP 0373335-A 33 20-JUN-1990;
              Gruenenthal GmbH
JOURNAL
FEATURES
source       Location/Qualifiers
              1. .321
              /organism="synthetic construct"
              /db_xref="taxon:32630"
gene         1. .321
              /gene="ALP-242"
CDS          1. .321
              /partial
              /gene="ALP-242"
              /codon_start=1
              /translation="SGKSEKAGVCPPKKSAQCLRYKKPEQSDWQCPGKRRCCPDTCG
              IKCLDPVDTNPTRRKPGKCPVTYGGCLMLNPPNFCEMDGQCKRDLKCCMGCMGKSCV
              SPVKA"
BASE COUNT   71 a   85 c   91 g   74 t
ORIGIN
Alignment Scores:
Pred. No.:      2.51e-07      Length:      321
Score:          140.00      Matches:      26
Percent Similarity: 55.36%      Conservative: 5
Best Local Similarity: 46.43%      Mismatches: 23
Query Match:     42.81%      Indels: 2
DB:              6      Gaps: 1

09-833799-13B (1-57) x A08111 (1-321)
QY      3 GluProValLysGlyProValSerThr-----LysProGlySerCysProIleIleLeu 20
      ::||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db      145 GATCCGGTTGACACCCGAAACCCGACGCGTCGTAAACCCGGGAAGTGCCCGTTACCTAC 204
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY      21 IleArgCysAlaMetLeuAsnProProAsnArgCysLeuLysAspThrAspCysProGly 40
      ::||| ::||| ::||| ::||| ::||| ::||| ::||| ::||| ::||| ::||| |||
Db      205 GGTCAAGTGCCTGCTGTAACCCGCTAACTTCTCGAGATGGATGGCCAGTGCAAAACGA 264
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY      41 IleLysLysCysCysGluGlySerCysGlyMetAlaCysPheValPro 56
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db      265 GATCTGAAATGCTGCATGGGTATGTGCGGTAAAGCTGCGTTAGCCCG 312
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
RESULT 67
A08112
LOCUS      A08112      321 bp      DNA      linear      PAT 28-JUL-1993
DEFINITION   Synthetic ALP-gene 246.
ACCESSION    A08112
VERSION      A08112.1  GI:413360
KEYWORDS
SOURCE       synthetic construct.
ORGANISM     synthetic construct
              artificial sequences.
REFERENCE    1 (bases 1 to 321)
AUTHORS      Heinzel-Wieland,R., Ammann,J., Steffens,G.J. and Flohe,L.
TITLE        Serine protease inhibitor proteins, medicaments containing them,
              DNA sequences coding for these proteins and methods for producing
              these proteins, medicaments and DNA sequences
              Patent: EP 0373335-A 34 20-JUN-1990;
              Gruenenthal GmbH
JOURNAL
FEATURES
source       Location/Qualifiers
```



```

VERSION      A31082.1  GI:1249296
KEYWORDS
SOURCE       synthetic construct.
ORGANISM     synthetic construct
              artificial sequences.
REFERENCE    1 (bases 1 to 74)
AUTHORS      Christophers,E., Schroder,J.M., Pioli,D., Wiedow,O. and Edge,M.D.
TITLE        Polypeptides and polypeptide analogues with inhibitory activity
              against human elastase
JOURNAL      Patent: EP 0402068-A 22 12-DEC-1990;
              IMPERIAL CHEMICAL INDUSTRIES PLC
FEATURES     Location/Qualifiers
              1..74
              /organism="synthetic construct"
              /db_xref="taxon:32630"
BASE COUNT   19 a      14 c      18 g      23 t
ORIGIN
Alignment Scores:
Pred. No.:      2e-07      Length:      74
Score:          135.00     Matches:    22
Percent Similarity: 100.00%  Conservative: 0
Best Local Similarity: 100.00%  Mismatches: 0
Query Match:    41.28%      Indels:     0
DB:              6          Gaps:         0

09-833799-13B (1-57) x A31082 (1-74)

QY  36  ThrAspCysProGlyIleLysLysCysCysGluGlySerCysGlyMetAlaCysPheVal 55
      |||||
Db   3  ACTGATTGTCAGGTATCAAAAGTGCTGTGAAGGTTTCCTGCGGTATGGCTTGTTCGTT 62

QY  56  ProGln 57
      |||||
Db   63  CCACAA 68

RESULT 72
A08109
LOCUS       A08109
DEFINITION  Synthetic ALP-gene 236.
ACCESSION   A08109
VERSION     A08109.1  GI:413354
KEYWORDS    .
SOURCE      synthetic construct.
            synthetic construct
            artificial sequences.
REFERENCE   1 (bases 1 to 321)
AUTHORS     Heinzel-Wieland,R., Ammann,J., Steffens,G.J. and Flohe,L.
TITLE       Serine protease inhibitor proteins, medicaments containing them,
            DNA sequences coding for these proteins and methods for producing
            these proteins, medicaments and DNA sequences
            Patent: EP 0373335-A 31 20-JUN-1990;
            Gruenthal GmbH
FEATURES     Location/Qualifiers
              1..321
              /organism="synthetic construct"
              /db_xref="taxon:32630"
              1..321
              /gene="ALP-236"
              1..321
              /partial
              /gene="ALP-236"
              /codon_start=1
              /transl_table=11
              /protein_id="CAA00745.1"
              /db_xref="GI:413355"
              /translation="SGKSFKAGVCPPKKSAQCLRYKKPEQCSDWQCPGKKRCCPDTCG
              IKCLDPVDTPNPTRRKP GKCPVTEGQCLLLNPPNFCELDQCKRDLKCCLGLCGKSCV
              SPVKA"
BASE COUNT   69 a      88 c      91 g      73 t
ORIGIN
Alignment Scores:

```

Pred. No.: 9.3e-07 Length: 321
Score: 135.00 Matches: 26
Percent Similarity: 55.36% Conservatives: 5
Best Local Similarity: 46.43% Mismatches: 23
Query Match: 41.28% Indels: 2
DB: 6 Gaps: 1

09-833799-13B (1-57) x A08109 (1-321)

QY 3 GluProValLysGlyProValSerThr-----LysProGlySerCysProIleIleLeu 20

Db 145 GATCCGGTTGACACCCCGAACCGACGCGTCGTTAAACCCGGGAAGTGCCCGTTACCGAA 204

QY 21 IleArgCysAlaMetLeuAsnProProAsnArgCysLeuLysAspThrAspCysProGly 40

Db 205 GGTCAGTGCCTGCTGTAACCGCGCTTAACCTTCGCGAGCTCGATGGCCAGTGCAAAACGA 264

QY 41 IleLysLysCysCysGluGlySerCysGlyMetAlaCysPheValPro 56

Db 265 GATCTGAAATGCTGCTGGGTCTGTGCGGTAAAGCTGCGTTAGCCCG 312

RESULT 73
AB003285
LOCUS AB003285 1054 bp DNA linear MAM 26-MAY-1999
DEFINITION Sus scrofa gene for elafin homolog, partial cds.
ACCESSION AB003285
VERSION AB003285.1 GI:4887645
KEYWORDS elafin homolog.
SOURCE Sus scrofa DNA.
ORGANISM Sus scrofa

REFERENCE
AUTHORS Furutani, Y., Kato, A., Yasue, H., Alexander, L.J., Beattie, C.W. and Hirose, S.
TITLE Evolution of the trappin multigene family in the Suidae
JOURNAL J. Biochem. 124 (3), 491-502 (1998)
MEDLINE 98391820
REFERENCE 2 (bases 1 to 1054)
AUTHORS Hirose, S.
TITLE Direct Submission
JOURNAL Submitted (21-APR-1997) Shigehisa Hirose, Tokyo Institute of Technology, Department of Biological Sciences; 4259 Nagatsuta-cho, Midori-ku, Yokohama, Kanagawa 226-8501, Japan
(E-mail: shirose@bio.titech.ac.jp, Tel: 045-924-5726, Fax: 045-924-5824)

REFERENCE
AUTHORS Furutani, Y., Kato, A., Yasue, H., Alexander, L.J., Beattie, C.W. and Hirose, S.
TITLE Evolution of the trappin multigene family in the Suidae
JOURNAL J. Biochem. 124 (3), 491-502 (1998)
MEDLINE 98391820
REFERENCE 2 (bases 1 to 1054)
AUTHORS Hirose, S.
TITLE Direct Submission
JOURNAL Submitted (21-APR-1997) Shigehisa Hirose, Tokyo Institute of Technology, Department of Biological Sciences; 4259 Nagatsuta-cho, Midori-ku, Yokohama, Kanagawa 226-8501, Japan
(E-mail: shirose@bio.titech.ac.jp, Tel: 045-924-5726, Fax: 045-924-5824)

FEATURES
source Location/Qualifiers
1..1054
/organism="Sus scrofa"
/db_xref="taxon:9823"
CDS <211..557
/codon_start=3
/product="elafin homolog"
/protein_id="BAA77829.1"
/db_xref="GI:4887645"
translation="GLKGGQDPVAGQAEQGVKVEILDIGQDLVKSQDPVEGQDPV
QAQLPDKVQDPVKAQPAVQGVFFSKLGFCEPRICRLNPNRCLIDAQCPGFQKCC
RVCGVKSCADPR"

exon 263 a 262 c 307 g 222 t
BASE COUNT 211..>557
ORIGIN
Alignment Scores:
Pred. No.: 3.24e-06 Length: 1054
Score: 135.00 Matches: 32
Percent Similarity: 49.40% Conservatives: 9
Best Local Similarity: 38.55% Mismatches: 14
Query Match: 41.28% Indels: 28
DB: 4 Gaps: 5

09-833799-13B (1-57) x AB003285 (1-1054)

QY 1 AlaGlnGluProValLysGly-----

Db 312 AGTCAAGATCCAGTTGAGGTCAAGATCCAGTCCAGCCCAACTTCCAGACAAAGTACAA 371

QY 8 ---ProValSerThrLysPro-----GlySer 15

Db 372 GATCCAGTCAAAGCCCACTGCAGTCCAAGGTGATTCCTTTCTCTAAGCTTGGCTTC 431

QY 16 CysProIleIleLeuIleArgCysAlaMetLeuAsnProProAsnArgCysLeuLysAsp 35

Db 432 TGCCCCAGGATTGAGATCCGTTGCAGACTGCTCAAT---CCTAACAGGTGTTTATAGAT 488

QY 36 ThrAspCysProGlyIleLysLysCysCysGluGlySerCysGlyMet---AlaCysPhe 54

Db 489 GCTCAGTCCCAGGGTTCAGAAAGTGTGTAGA---GTCTGTGGTGTGAAGTCTCTGCG 545

QY 55 ValProGln 57

Db 546 GATCCCCGG 554

RESULT 74
A08106
LOCUS A08106 Synthetic ALP-gene 240.
DEFINITION A08106
ACCESSION A08106
VERSION A08106.1 GI:413348
KEYWORDS synthetic construct.
SOURCE synthetic construct.
ORGANISM artificial construct.
REFERENCE 1 (bases 1 to 321)
AUTHORS Heinzel-Wieland, R., Ammann, J., Steffens, G.J. and Flohe, L.
TITLE Serine protease inhibitor proteins, medicaments containing them, DNA sequences coding for these proteins and methods for producing these proteins, medicaments and DNA sequences
Patent: EP 0373335-A 28 20-JUN-1990;
Gruenthal GmbH

JOURNAL Location/Qualifiers
1..321
/organism="synthetic construct"
/db_xref="taxon:32630"
gene 1..321
/gene="ALP-240"
CDS 1..321
/partial
/gene="ALP-240"
/codon_start=1
/transl_table=11
/protein_id="CAA00742.1"
/db_xref="GI:413348"
/translation="SGKSFKAGVCPPKSAQCLRYKKPEQSDWQCQPKKRCPPDTCC
IKCLDPVDTNPTRRRKPGKCPVYQGCLLLNPNFCEMDGQCKRDLKCCMGCKSCV
SPVKA"

BASE COUNT 72 a 83 c 92 g 74 t
ORIGIN

Alignment Scores:
Pred. No.: 2.65e-06 Length: 321
Score: 131.00 Matches: 25
Percent Similarity: 53.57% Conservatives: 5
Best Local Similarity: 44.64% Mismatches: 24
Query Match: 40.06% Indels: 2
DB: 6 Gaps: 1

09-833799-13B (1-57) x A08106 (1-321)

QY 3 GluProValLysGlyProValSerThr-----LysProGlySerCysProIleIleLeu 20

Db 145 GATCCGGTTGACACCCCGAACCGACGCGTCGTAAACCCGGGAAGTGCCCGTTACCTAC 204

QY 21 IleArgCysAlaMetLeuAsnProProAsnArgCysLeuLysAspThrAspCysProGly 40

Db 205 GGTCAGTGCCTGCTGCTGAACAGGCGCTAACTTCTCGAGATGGATGGCCAGTGCAAACGA 264

GenCore version 5.1.4_p5_4578
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: February 15, 2003, 19:00:51 ; Search time 223 Seconds
(without alignments)
575.623 Million cell updates/sec

Title: 09-833799-13B
Perfect score: 327
Sequence: 1 aqepvkvstkpqscpiil.....cpgikkcccgscgmcfvpq 57

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 2185239 seqs, 1125999159 residues

Total number of hits satisfying chosen parameters: 4370478

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 100 summaries

Command line parameters:
-MODEL=frame+p2n.model -DEV=xlp
-Q=/cgn2_1/USPTO_spool/US09833799/runat_12022003_120041_23624/app_query.fasta_1.199
-DB=N_Geneseq_101002 -QFMT=fastap -SUFFIX=rng -MINMATCH=0.1 -LOOPCL=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=-1 -MATRIX=blosum62 -TRANS=human40.cdi
-LIST=100 -DOCALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=75
-MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US09833799@cgn_1_1_396 @runat_12022003_120041_23624 -NCPU=6 -ICPU=3
-NO_XLPXY -NO_MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -LONGLOG -DEV_TIMEOUT=120
-WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : N_Geneseq_101002: *
1: /SIDS2/gcgdata/geneseq/geneseq-embl/NA1980.DAT: *
2: /SIDS2/gcgdata/geneseq/geneseq-embl/NA1981.DAT: *
3: /SIDS2/gcgdata/geneseq/geneseq-embl/NA1982.DAT: *
4: /SIDS2/gcgdata/geneseq/geneseq-embl/NA1983.DAT: *
5: /SIDS2/gcgdata/geneseq/geneseq-embl/NA1984.DAT: *
6: /SIDS2/gcgdata/geneseq/geneseq-embl/NA1985.DAT: *
7: /SIDS2/gcgdata/geneseq/geneseq-embl/NA1986.DAT: *
8: /SIDS2/gcgdata/geneseq/geneseq-embl/NA1987.DAT: *
9: /SIDS2/gcgdata/geneseq/geneseq-embl/NA1988.DAT: *
10: /SIDS2/gcgdata/geneseq/geneseq-embl/NA1989.DAT: *
11: /SIDS2/gcgdata/geneseq/geneseq-embl/NA1990.DAT: *
12: /SIDS2/gcgdata/geneseq/geneseq-embl/NA1991.DAT: *
13: /SIDS2/gcgdata/geneseq/geneseq-embl/NA1992.DAT: *
14: /SIDS2/gcgdata/geneseq/geneseq-embl/NA1993.DAT: *
15: /SIDS2/gcgdata/geneseq/geneseq-embl/NA1994.DAT: *
16: /SIDS2/gcgdata/geneseq/geneseq-embl/NA1995.DAT: *
17: /SIDS2/gcgdata/geneseq/geneseq-embl/NA1996.DAT: *
18: /SIDS2/gcgdata/geneseq/geneseq-embl/NA1997.DAT: *
19: /SIDS2/gcgdata/geneseq/geneseq-embl/NA1998.DAT: *
20: /SIDS2/gcgdata/geneseq/geneseq-embl/NA1999.DAT: *
21: /SIDS2/gcgdata/geneseq/geneseq-embl/NA2000.DAT: *
22: /SIDS2/gcgdata/geneseq/geneseq-embl/NA2001A.DAT: *
23: /SIDS2/gcgdata/geneseq/geneseq-embl/NA2001B.DAT: *
24: /SIDS2/gcgdata/geneseq/geneseq-embl/NA2002.DAT: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB	ID	Description
1	327	100.0	206	11	AAQ06819	Sequence encoding
c 2	327	100.0	434	21	AAC68808	Human head/neck tu
c 3	327	100.0	480	24	ABK45842	CDNA encoding colo
4	327	100.0	504	11	AAQ06820	Sequence encoding
5	327	100.0	571	24	ABQ58818	Human colon cancer
6	327	100.0	2309	24	ABK83819	Human CDNA differe
7	327	100.0	2309	24	ABL65816	Lung cancer relate
8	327	100.0	2309	24	ABL66488	Lung cancer relate
9	324	99.1	177	15	AAQ56674	Elafin derivative
10	324	99.1	8598	19	AAV28852	ppIC9/ELF25L cDNA
11	323	98.8	177	15	AAQ56675	Elafin derivative
12	323	98.8	177	15	AAQ56676	Elafin derivative
13	322	98.5	321	13	AAQ28757	Partial sequence o
14	179.5	54.9	737	15	AAQ44862	SPAI gene. Sus sc
15	161	49.2	396	24	ABQ73674	Murine SLPI sense
c 16	161	49.2	396	24	ABQ73677	Murine SLPI antise
17	161	49.2	684	20	AAI18516	Mouse IMC carcinom
18	161	49.2	691	20	AAI18514	Mouse IMC carcinom
19	161	49.2	1114	20	AAI18515	Mouse IMC carcinom
20	146	44.6	180	21	AAC97526	DNA encoding a ser
21	146	44.6	183	20	AAI16273	Secretory leukocyt
22	146	44.6	183	22	AAI67584	Secretory leukocyt
23	143	43.7	194	10	AAN90354	Sequence encoding
24	143	43.7	321	24	ABK8016	DNA encoding human
25	143	43.7	324	7	AAN60463	Synthetic sequence
26	143	43.7	324	7	AAN60464	Synthetic sequence
27	143	43.7	324	20	AAI16194	Serine protease in
28	143	43.7	324	20	AAI16236	Serine protease in
29	143	43.7	324	21	AAC97527	DNA encoding a ser
30	143	43.7	324	21	AAC97528	Secretory leukocyt
31	143	43.7	324	21	AAC97579	DNA encoding a ser
32	143	43.7	324	22	AAI67488	Recombinant serine
33	143	43.7	324	22	AAI67489	Recombinant secret
34	143	43.7	399	22	ABA83108	Secretory leukocyt
35	143	43.7	420	20	AAI16272	DNA sequence of om
36	143	43.7	436	21	AAC97624	DNA encoding ompA-
37	143	43.7	436	22	AAI67583	DNA sequence codin
38	143	43.7	460	20	AAI16271	DNA sequence of om
39	143	43.7	460	21	AAC97622	DNA encoding OmpA
40	143	43.7	460	22	AAI67582	DNA sequence codin
41	143	43.7	498	8	AAN70929	Sequence encoding
42	143	43.7	565	8	AAN70930	Sequence encoding
43	143	43.7	565	22	AAF32043	Human antileukopro
44	143	43.7	594	24	ABL66679	Lung cancer relate
45	143	43.7	594	24	ABL67548	Thyroid cancer rel
46	143	43.7	594	24	AAI17499	Human cDNA encodin
47	143	43.7	599	20	AAZ77546	Human ovarian tumo
48	143	43.7	636	22	AAH57451	Human lung cell sp
49	143	43.7	1525	24	ABK88022	DNA sequence encod
50	143	43.7	1525	24	ABK88025	DNA sequence encod
51	141	43.1	2274	7	AAN60465	Synthetic sequence
52	141	43.1	2274	21	AAC97530	Human genomic DNA
53	141	43.1	2274	22	AAI67491	Recombinant secret
54	139	42.5	180	15	AAQ45442	CLPI. Homo sapien
55	139	42.5	180	21	AAZ88483	Truncated SLPI enc
56	123.5	37.8	1010	21	AAZ37665	Human peptidase, H
57	123.5	37.8	1013	24	AAI38698	Human LP229 secret
58	122	37.3	321	21	AAC97529	DNA encoding a pro
59	121.5	37.2	1411	21	AAZ51673	Human p53 target m
60	118.5	36.2	558	22	AAH78211	Nucleotide sequenc
61	118.5	36.2	732	22	AAH78210	Nucleotide sequenc
62	118.5	36.2	5690	22	AAI07251	Human reproductive
63	118.5	36.2	5690	23	ABL98799	Human testicular a
64	116	35.5	872	7	AAN60469	ompA-tc-met-secret
65	116	35.5	920	7	AAN60468	ompA-secretory leu
66	112	34.3	478	20	AAH80909	Human cDNA clone H
67	112	34.3	762	21	AAA99904	CDNA encoding huma

68 112 34.3 762 21 AAZ65078 Membrane-bound pro
69 112 34.3 762 22 AAF44224 Human PRO844 (UNQ5
70 107 32.7 411 24 ABK64114 Lung benign prost
71 107 32.7 411 24 ABL66677 Lung cancer relate
72 107 32.7 411 24 ABL67461 Thyroid cancer rel
73 107 32.7 411 24 ABL67811 Ovary cancer relat
74 105 32.1 292 21 AAA70066 Human ovarian carc
75 105 32.1 292 24 ABN72960 Ovarian carcinoma
76 103 31.5 4536 23 ABL09753 Drosophila melanog
77 103 31.5 8853 23 ABL09752 Drosophila melanog
78 101 30.9 1796 23 ABL10651 Drosophila melanog
79 101 30.9 7930 23 ABL10650 Drosophila melanog
80 98 30.0 705 22 AAF59010 Mouse epididymis-s
81 98 30.0 724 22 AAD10127 Mouse major epidid
82 98 30.0 724 24 ABL35060 Murine cDNA isolat
83 98 30.0 767 24 AAD37774 Extended sequence
84 98 30.0 8577 23 ABL28667 Drosophila melanog
85 98 30.0 9295 23 ABL02495 Drosophila melanog
86 98 30.0 12294 23 ABL28666 Drosophila melanog
87 98 30.0 21314 23 ABL02494 Drosophila melanog
88 97 29.7 120 21 AAC97577 Insert C used in p
89 97 29.7 120 21 AAC97620 DNA insert used in
90 97 29.7 120 22 AAI67539 Serine protease in
91 97 29.7 583 12 AAQ13127 HE4 epididymis-spe
92 97 29.7 583 19 AAV64622 Human HE4 cDNA. H
93 97 29.7 583 22 AAD02714 Human epididymal p
94 97 29.7 583 24 ABL68022 Ovary cancer relat
95 96.5 29.5 308 24 ABN25109 Human ORFX polynuc
96 96.5 29.5 401 22 AAK54133 Murine transport a
97 95.5 29.2 500 21 AAC93641 Cat flea head and
98 95 29.1 469 22 ABA42966 Human breast cell
99 95 29.1 469 22 ABA23163 Probe #1629 for ge
100 95 29.1 469 22 AAK01653 Human brain expres

ALIGNMENTS

RESULT 1
AAQ06819
ID AAQ06819 standard; DNA; 206 BP.

AC AAQ06819;
XX
DT 06-MAR-1991 (first entry)
XX
DE Sequence encoding human leukocyte elastase inhibiting peptide.
XX
KW Emphysema; respiratory distress syndrome; atherosclerosis;
KW arthritis; cystic fibrosis; leukaemia.
XX
OS Homo sapiens.

Key Location/Qualifiers
CDS 30..203
FT /*tag= a
FT /label= Elastase inhibitor

XX EP402068-A.
XX
PD 12-DEC-1990.
XX
PF 04-JUN-1990; 90EP-0306037.
XX
PR 02-NOV-1989; 89GB-0024717.
PR 09-JUN-1989; 89GB-0013346.
PR 09-JUN-1989; 89GB-0013349.
PR 25-SEP-1989; 89GB-0021613.

XX (ICIL) IMPERIAL CHEM INDS PLC.
PA
XX Christophers E, Schroder JM, Pioli D, Wiedow O, Edge MD;
PI
XX WPI; 1990-370081/50.
DR

DR P-PSDB; AAR08217.
XX Human leukocyte elastase inhibitor - is genetically engineered
PT polypeptide for treatment of inflammatory, pulmonary and skin
PT conditions
XX
PS Disclosure; Fig 13; 45pp; English.
XX
CC Gene product has inhibitory activity against human leukocyte
CC elastase (HLE) and may be isolated from psoriatic scales. The
CC sequence may be expressed from a plasmid transformed expression
CC system and may be useful in the prevention of tissue damage
CC associated with emphysema, adult respiratory distress syndrome,
CC psoriasis and bullous dermatoses. Other treatable conditions
CC include atherosclerosis, cystic fibrosis, bronchitis and acute
CC non-lymphoblastic leukaemia.
CC Abs raised to the polypeptide may be used in detection.
XX
SQ Sequence 206 BP; 51 A; 46 C; 44 G; 65 T; 0 other;

Alignment Scores:
Pred. No.: 2,47e-29 Length: 206
Score: 327.00 Matches: 57
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 11 Gaps: 0

09-833799-13B (1-57) x AAQ06819 (1-206)

Qy 1 AlaGlnGluProValLysGlyProValSerThrLysProGlySerCysProIleLeu 20
Db 30 GCTCAAGAACCAAGTAAAGGTCCTGTGTCTACTAAGCCAGGTTCTTGCTTATCTTG 89

Qy 21 IleArgCysAlaMetLeuAsnProProAsnArgCysLeuLysAspThrAspCysProGly 40
Db 90 ATTCGTGCGGTATGTTAAACCCACCTAACCGTTGTTGAAGGACACTGATTGTCCAGGT 149
Qy 41 IleLysLysCysCysGluGlySerCysGlyMetAlaCysPheValProGln 57
Db 150 ATCAAAAAGTGCTGTGAAGTTCCTGCGGTATGCGTGTGTTTCGTTCCACAA 200

RESULT 2
AAC68808/c
ID AAC68808 standard; cDNA; 434 BP.
XX
AC AAC68808;
XX
DT 21-FEB-2001 (first entry)
XX
DE Human head/neck tumour related protein partial coding sequence #5.
XX
KW Head tumour; neck tumour; lung cancer; vaccine; cancer therapy; ss.
XX
OS Homo sapiens.
XX
PN WO200065053-A2.
XX
PD 02-NOV-2000.
XX
PF 21-APR-2000; 2000WO-US10687.
XX
PR 23-APR-1999; 99US-0130906.
PR 20-APR-2000; 2000US-0533870.
XX
PA (CORI-) CORIXA CORP.
XX
PI Wang T, Dillon DC;
XX
DR WPI; 2000-687345/67.

XX Novel polypeptides comprising immunogenic portion of head and neck
PT tumour protein useful for treating, diagnosing and monitoring cancer

PT such as head, neck and lung cancer -
XX
PS Claim 3; Page 71; 77pp; English.
XX

CC The present invention relates to a number of nucleic acid sequences which
CC encode proteins associated with head, neck and lung tumours. These
CC tumours are often not diagnosed until they have spread, and, of those who
CC survive, most must endure alterations in facial and neck appearance as
CC well as changes in speech, sight, smell, chewing, swallowing and taste
CC perception. The coding sequences given, and the proteins they encode, can
CC be used in the diagnosis, treatment and vaccination against cancer,
CC particularly papillary and follicular carcinomas, papillary tumours,
CC follicular adenoma, parathyroid hyperplasia, parotid cancer, lip cancer,
CC squamous cell cancer of the tongue, oral tongue cancers and larynx
CC cancer.

XX
SQ Sequence 434 BP; 103 A; 103 C; 133 G; 94 T; 1 other;

Alignment Scores:
Pred. No.: 6.15e-29 Length: 434
Score: 327.00 Matches: 57
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 21 Gaps: 0

09-8333799-13B (1-57) x AAC68808 (1-434)

QY 1 AlaGlnGluProValLysGlyProValSerThrLysProGlySerCysProIleLeu 20
Db 318 GCGAAGAGCCAGTCAAGGTCAGTCTCCACTAAGCCTGGCTCCTGCCCATTTATCTTG 259
QY 21 IleArgCysAlaMetLeuAsnProProAsnArgCysLeuLysAspThrAspCysProGly 40
Db 258 ATCCGGTGCAGCATGTTGAATCCCCCTAACCCGCTGCTTGAAAGATACACTGACTGCCCCAGGA 199
QY 41 IleLysLysCysCysGluGlySerCysGlyMetAlaCysPheValProGln 57
Db 198 ATCAAGAAGTGCTGTGAAGGCTCTTGCGGGATGGCCTGTTTCGTTCCCCAG 148

RESULT 3
ID ABK45842/c
XX ABK45842 standard; cDNA; 480 BP.
AC ABK45842;
XX
DT 05-JUN-2002 (first entry)
DE cDNA encoding colon tumour protein, SEQ ID No 1393.
XX
KW Human; colon tumour; vaccine; colon cancer; immunogenic;
XX immunotherapy; gene; ss.
OS Homo sapiens.
XX
PN WO200212328-A2.
XX
PD 14-FEB-2002.
XX
PF 31-JUL-2001; 2001WO-US24218.
XX
PR 03-AUG-2000; 2000US-223283P.
PR 28-MAR-2001; 2001US-279763P.
PR 29-JUN-2001; 2001US-302051P.
XX
PA (CORI-) CORIXA CORP.
XX
PI King GE, Meagher MJ, Xu J, Secrist H;
XX
DR WPI; 2002-241739/29.
XX

PT New colon cancer polypeptides and polynucleotides, useful as vaccines,
PT for diagnosing, preventing, and treating colon cancer, and as markers

PT for the progression of cancer -
XX
PS Claim 1; SEQ ID No 1393; 147pp; English.
XX

CC The invention relates to polynucleotides encoding colon tumour proteins.
CC The polynucleotides and encoded polypeptides are useful in pharmaceutical
CC compositions, such as vaccines, for the diagnosis, prevention, and
CC treatment of colon cancer. Polynucleotide sequences may be used as
CC hybridisation probes or primers, and in the design and preparation of
CC ribozyme molecules for inhibiting expression of tumour polypeptides and
CC proteins in tumour cells. The compositions are useful for stimulating an
CC immune response against cancer, particularly for the immunotherapy of
CC colon cancer, and as markers for the progression of cancer.
CC ABK44450-ABK46237 represent coding sequences of human colon tumour
CC proteins of the invention.

CC Note: With the exception of SEQ ID No 1 and 2, the sequence data
CC for this patent did not form part of the printed specification but was
CC supplied by the European Patent Office.

XX
SQ Sequence 480 BP; 113 A; 105 C; 139 G; 120 T; 3 other;

Alignment Scores:
Pred. No.: 6.96e-29 Length: 480
Score: 327.00 Matches: 57
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 24 Gaps: 0

09-8333799-13B (1-57) x ABK45842 (1-480)

QY 1 AlaGlnGluProValLysGlyProValSerThrLysProGlySerCysProIleLeu 20
Db 369 GCGAAGAGCCAGTCAAGGTCAGTNTCCACTAAGCCTGGCTCCTGCCCATTTATCTTG 310
QY 21 IleArgCysAlaMetLeuAsnProProAsnArgCysLeuLysAspThrAspCysProGly 40
Db 309 ATCCGGTGCAGCATGTTGAATCCCCCTAACCCGCTGCTTGAAAGATACACTGACTGCCCCAGGA 250
QY 41 IleLysLysCysCysGluGlySerCysGlyMetAlaCysPheValProGln 57
Db 249 ATCAAGAAGTGCTGTGAAGGCTCTTGCGGGATGGCCTGTTTCGTTCCCCAG 199

RESULT 4
AAQ06820
ID AAQ06820 standard; cDNA; 504 BP.
XX
AC AAQ06820;
XX
DT 06-MAR-1991 (first entry)
DE
XX Sequence encoding human leukocyte elastase inhibiting peptide.
DE
XX HLE; Emphysema; respiratory distress syndrome; atherosclerosis;
KW arthritis; cystic fibrosis; leukaemia; ds.
XX
OS Homo sapiens.
XX

FH Key Location/Qualifiers
FT CDS 1..157
FT /*tag= b
FT /label= Upstream in-frame coding sequence
FT CDS 121..328
FT /*tag= a
FT /label= Elastase inhibitor
FT intron 144..180
FT /*tag= c
XX

PN EP402068-A.
XX
PD 12-DEC-1990.
XX
PF 04-JUN-1990; 90EP-0306037.

XX 02-NOV-1989; 89GB-0024717.
PR 09-JUN-1989; 89GB-0013346.
PR 09-JUN-1989; 89GB-0013349.
PR 25-SEP-1989; 89GB-0021613.
XX (ICIL) IMPERIAL CHEM INDS PLC.
PA Christophers E, Schroder JM, Pioli D, Wiedow O, Edge MD;
XX
XX WPI; 1990-370081/50.
DR P-PSDB; AAR08217.
DR
XX Human leucocyte elastase inhibitor - is genetically engineered
PT polypeptide for treatment of inflammatory, pulmonary and skin
PT conditions
PT
XX Disclosure; Fig 16; 45pp; English.
PS
XX Gene product has inhibitory activity against human leukocyte
CC elastase (HLE) and may be isolated from psoriatic scales. The
CC sequence may be expressed from a plasmid transformed expression
CC system and may be useful in the prevention of tissue damage
CC associated with emphysema, adult respiratory distress syndrome,
CC psoriasis and bullous dermatoses. Other treatable conditions
CC include atherosclerosis, cystic fibrosis, bronchitis and acute
CC non-lymphoblastic leukaemia.
CC Abs raised to the polypeptide may be used in detection.
XX
SQ Sequence 504 BP; 110 A; 152 C; 120 G; 122 T; 0 other;

Alignment Scores:
Pred. No.: 7.39e-29 Length: 504
Score: 327.00 Matches: 57
Percent Similarity: 100.00% Conservatives: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 11 Gaps: 0

09-833799-13B (1-57) x AAQ06820 (1-504)

QY 1 AlaGlnGluProValLysGlyProValSerThrLysProGlySerCysProIleLeu 20
Db 158 GCGCAAGAGCCAGTCAAGGTCCTCCACTAAGCCTGGCTCCTGCCCATATCTTG 217

QY 21 IleArgCysAlaMetLeuAsnProAsnArgCysLeuLysAspThrAspCysProGly 40
Db 218 ATCCGGTGGCCATGTTGAATCCCTTAACCGCTGCTTGAAGATACTGACTGCCAGGA 277

QY 41 IleLysLysCysCysGluGlySerCysGlyMetAlaCysPheValProGln 57
Db 278 ATCAAGAAGTGCTGTGAAGGCTCTTGGGGATGGCCTGTTTCGTTCCCCAG 328

RESULT 5
ABQ58818
ID ABQ58818 standard; cDNA; 571 BP.
XX
AC ABQ58818;
XX
DT 02-AUG-2002 (first entry)
XX
DE Human colon cancer related nucleotide sequence SEQ ID NO:2513.
XX
KW Human; colon cancer; cancer; tissue profiling; forensic; mapping;
KW genetic analysis; diagnostic; antisense therapy; gene; ss.
XX
OS Homo sapiens.
XX
PN WO200229086-A2.
XX
PD 11-APR-2002.
XX
PF '02-OCT-2001; 2001WO-US30732.

XX 02-OCT-2000; 2000US-237271P.
PR (FARB) BAYER CORP.
XX
XX Burgess C, Astle JH, Carroll E, Catino TJ, Dwivedi P, Molino GA;
PI Thiagalingam A, Lewis ME;
XX
XX WPI; 2002-426115/45.
XX
XX New isolated nucleic acid that is differentially expressed in cancer
PT tissues useful for determining the presence of colon cancer in a cell
PT or tissue type, and in antisense therapy -
XX
PS Claim 1; Fig 1; 796pp; English.
XX
CC ABQ56306 to ABQ60787 represent isolated nucleic acids (I) differentially
CC expressed in cancer tissues. ABB78993 to ABB79004 represent proteins
CC encoded by the ABQ60776 to ABQ60787 nucleic acid sequences. (I) can be
CC used in antisense therapy. An antibody immunoreactive with a polypeptide
CC encoded by (I) is useful for detecting cancer in a patient sample, and
CC for detecting the presence or absence of a polynucleotide encoded by a
CC nucleic acid which hybridizes to (I) in a cell. A probe/primer derived
CC from (I) can be used for determining the presence of a nucleic acid which
CC hybridizes to (I), and for determining the phenotype of cells in a sample
CC of cells from a patient. (I) is useful for determining the presence of
CC colon cancer in a cell or tissue type, for determining the presence or
CC state of other type of cancer, in antisense therapy, to generate
CC macroarrays on a solid surface, to identify a chromosome on which the
CC corresponding gene resides, and in tissue profiling, forensics, genetic
CC analysis, mapping and diagnostic applications. (I) can be used to raise
CC antibodies, and to screen for peptide analogues and antagonists.
XX
SQ Sequence 571 BP; 136 A; 155 C; 131 G; 145 T; 4 other;

Alignment Scores:
Pred. No.: 8.61e-29 Length: 571
Score: 327.00 Matches: 57
Percent Similarity: 100.00% Conservatives: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 24 Gaps: 0

09-833799-13B (1-57) x ABQ58818 (1-571)

QY 1 AlaGlnGluProValLysGlyProValSerThrLysProGlySerCysProIleLeu 20
Db 210 GCGCAAGAGCCAGTCAAGGTCCTCCACTAAGCCTGGCTCCTGCCCATATCTTG 269

QY 21 IleArgCysAlaMetLeuAsnProAsnArgCysLeuLysAspThrAspCysProGly 40
Db 270 ATCCGGTGGCCATGTTGAATCCCTTAACCGCTGCTTGAAGATACTGACTGCCAGGA 329

QY 41 IleLysLysCysCysGluGlySerCysGlyMetAlaCysPheValProGln 57
Db 330 ATCAAGAAGTGCTGTGAAGGCTCTTGGGGATGGCCTGTTTCGTTCCCCAG 380

RESULT 6
ABK83819
ID ABK83819 standard; cDNA; 2309 BP.
XX
AC ABK83819;
XX
DT 14-AUG-2002 (first entry)
XX
DE Human cDNA differentially expressed in granulocytic cells #390.
XX
KW Human; ss; granulocytic cell; DNA chip; bacterial infection;
KW viral infection; parasitic infection; protozoal infection;
KW fungal infection; sterile inflammatory disease; psoriasis;
KW rheumatoid arthritis; glomerulonephritis; asthma; thrombosis;
KW cardiac reperfusion injury; renal reperfusion injury; ARDS;
KW adult respiratory distress syndrome; inflammatory bowel disease;

Crohn's disease; ulcerative colitis; periodontal disease; granulocyte activation; chronic inflammation; allergy.

Homo sapiens.

WO200228999-A2.

11-APR-2002.

03-OCT-2001; 2001WO-US30821.

03-OCT-2000; 2000US-237189P.

(GENE-) GENE LOGIC INC.

Beazer-Barclay Y, Weissman SM, Yamaga S, Vockley J;

WPI; 2002-435328/46.

Detecting granulocyte activation by detecting differential expression of genes associated with granulocyte activation, which serves as diagnostic markers that is useful for monitoring disease states and drug toxicity -

Claim 1; SEQ ID No 390; 114pp; English.

The invention relates to detecting (M1) granulocyte (GC) activation (GCA), by detecting the level of expression of gene(s) (Gs) identified by DNA chip analysis as given in the specification, and comparing the expression level to an expression level in an unactivated GC, where differential expression of Gs is indicative of GCA.

Also included are modulating (M2) GA by contacting GC with an agent that alters the expression of at least one gene in Gs; (2) screening (M3) for an agent capable of modulating GCA or an inflammation (especially chronic) in a tissue, an allergic response in a subject, exposure of a subject to a pathogen or sterile inflammatory disease using the gene expression profile; (3) detecting (M4) an inflammation (especially chronic) in a tissue, an allergic response in a subject, exposure of a subject to a pathogen or sterile inflammatory disease, by detecting the level of expression in a sample of the tissue of gene(s) from Gs, where the level of expression of the gene is indicative of inflammation;

(4) treating (M5) an inflammation (especially chronic) or in a tissue, an allergic response in a subject, exposure of a subject to a pathogen or sterile inflammatory disease, by contacting a tissue having inflammation with an agent that modulates the expression of gene(s) from Gs in the tissue. M1 is useful for detecting GCA; M2 is useful for modulating GA; M3 is useful for screening an agent capable of modulating GCA preferably in an inflammation in a tissue; M4 is useful for detecting an inflammation (especially chronic) in a tissue, an allergic response in a subject, exposure of a subject to a pathogen or sterile inflammatory disease (e.g. psoriasis, rheumatoid arthritis, glomerulonephritis, asthma, thrombosis, cardiac reperfusion injury, renal reperfusion injury, ARDS, adult respiratory distress syndrome, inflammatory bowel disease, Crohn's disease, ulcerative colitis, periodontal disease; also bacterial infection, viral infection, parasitic infection, protozoal infection, fungal infection and M5 is useful for treating one of the above conditions. The present sequence represents a gene differentially expressed in granulocytes.

Note: The sequence data for this patent did not form part of the printed specification, but was obtained in electronic format directly from WIPO at

ftp.wipo.int/pub/published pct sequences.

Sequence 2309 BP; 595 A; 537 C; 603 G; 572 T; 2 other;

Alignment Scores:

Pred. No.:	4.77e-28	Length:	2309
Score:	327.00	Matches:	57
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	100.00%	Indels:	0
DB:	24	Gaps:	0

09-833799-13B (1-57) x ABK83819 (1-2309)

Qy 1 AlaGlnGluProValLysGlyProValSerThrLysProGlySerCysProIleLeu 20
 Db 1554 GCGCAAGAGCCAGTCAAAGGTCCAGTCTCCACTAAGCCTGGCTCCTGCCCATTTATCTTG 1613

Qy 21 IleArgCysAlaMetLeuAsnProProAsnArgCysLeuLysAspThrAspCysProGly 40
 Db 1614 ATCCGGTGGCCCATGTTGAATCCCCCTAACCGCTGCTTGAAAGATACTGACTGCCCAGGA 1673

Qy 41 IleLysLysCysCysGluGlySerCysGlyMetAlaCysPheValProGln 57
 Db 1674 ATCAAGAAGTGCTGTGAAGGCTCTTGCGGGGATGGCCTGTTTCGTTCCCCAG 1724

RESULT 7
 ABL65816
 ID ABL65816 standard; DNA; 2309 BP.
 XX
 AC ABL65816;
 XX
 DT 15-MAY-2002 (first entry)
 XX
 DE Lung cancer related gene sequence SEQ ID NO:4153.
 XX
 KW Human; cancer; colon; breast; ovary; oesophagus; kidney; thyroid;
 KW stomach; lung; prostate; pancreas; carcinoma; antitumour; cancerous;
 KW cytostatic; gene therapy; antineoplastic; Wilm's tumour; adenocarcinoma;
 KW gene; ds.
 XX
 OS Homo sapiens.
 XX
 PN WO200194629-A2.
 XX
 PD 13-DEC-2001.
 XX
 PF 30-MAY-2001; 2001WO-US10838.
 XX
 PR 05-JUN-2000; 2000US-209473P.
 PR 05-JUN-2000; 2000US-209531P.
 PR 18-SEP-2000; 2000US-233133P.
 PR 18-SEP-2000; 2000US-233617P.
 PR 20-SEP-2000; 2000US-234009P.
 PR 20-SEP-2000; 2000US-234034P.
 PR 20-SEP-2000; 2000US-234052P.
 PR 22-SEP-2000; 2000US-234509P.
 PR 22-SEP-2000; 2000US-234567P.
 PR 25-SEP-2000; 2000US-234923P.
 PR 25-SEP-2000; 2000US-234924P.
 PR 25-SEP-2000; 2000US-235077P.
 PR 25-SEP-2000; 2000US-235082P.
 PR 25-SEP-2000; 2000US-235134P.
 PR 25-SEP-2000; 2000US-235280P.
 PR 26-SEP-2000; 2000US-235637P.
 PR 26-SEP-2000; 2000US-235638P.
 PR 27-SEP-2000; 2000US-235711P.
 PR 27-SEP-2000; 2000US-235720P.
 PR 27-SEP-2000; 2000US-235840P.
 PR 27-SEP-2000; 2000US-235863P.
 PR 28-SEP-2000; 2000US-236028P.
 PR 28-SEP-2000; 2000US-236032P.
 PR 28-SEP-2000; 2000US-236033P.
 PR 28-SEP-2000; 2000US-236034P.
 PR 28-SEP-2000; 2000US-236109P.
 PR 28-SEP-2000; 2000US-236111P.
 PR 29-SEP-2000; 2000US-236842P.
 PR 29-SEP-2000; 2000US-236891P.
 PR 02-OCT-2000; 2000US-237172P.
 PR 02-OCT-2000; 2000US-237173P.
 PR 02-OCT-2000; 2000US-237278P.
 PR 02-OCT-2000; 2000US-237294P.
 PR 02-OCT-2000; 2000US-237295P.
 PR 02-OCT-2000; 2000US-237316P.
 PR 03-OCT-2000; 2000US-237425P.

PR 03-OCT-2000; 2000US-237598P.
PR 03-OCT-2000; 2000US-237604P.
PR 03-OCT-2000; 2000US-237606P.
PR 03-OCT-2000; 2000US-237608P.
PR 01-NOV-2000; 2000US-244867P.
PR 01-NOV-2000; 2000US-245084P.
XX
PA (AVAL-) AVALON PHARM.
XX
PI Young PE, Augustus M, Carter KC, Ebner R, Endress G, Horrigan S;
PI Soppet DR, Weaver Z;
XX
DR WPI; 2002-188264/24.
XX
XX Screening for anti-neoplastic agent involves exposing cells to a
PT chemical agent to be tested for anti-neoplastic activity, and
PT determining a change in expression of a gene of a signature gene set -
XX
PS Claim 1; SEQ ID 4153; 44pp; English.
XX
CC The present invention describes a method (M1) for screening for an
CC anti-neoplastic agent. The method involves exposing cells to a chemical
CC agent to be tested for anti-neoplastic activity, determining a change in
CC expression of at least one gene (I) of a signature gene set, where (I)
CC comprises a sequence (S) selected from 8447 sequences (given in ABL61664
CC to ABL70110), or is at least 95% identical to (S), where a change in
CC expression is indicative of anti-neoplastic activity. (I) has cytostatic
CC activity and can be used in gene therapy. M1 can be used for screening
CC an anti-neoplastic agent, and can be used for producing a product which
CC is the data collected with respect to the anti-neoplastic agent as a
CC result of M1, and the data is sufficient to convey the chemical
CC structure and/or properties of the agent. M1 can be used in the
CC treatment of cancer such as colon, breast, stomach, lung, thyroid,
CC oesophageal, ovarian, kidney, prostate or pancreatic cancer,
CC adenocarcinoma, carcinoma, clear cell cancer, infiltrating ductal cancer,
CC infiltrating lobular cancer, squamous cell carcinoma, neuroendocrine
CC carcinoma, papillary carcinoma and Wilm's tumour.
XX
SQ Sequence 2309 BP; 595 A; 537 C; 603 G; 572 T; 2 other;

Alignment Scores:
Pred. No.: 4.77e-28 Length: 2309
Score: 327.00 Matches: 57
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 24 Gaps: 0

09-833799-13B (1-57) x ABL65816 (1-2309)

Qy 1 AlaGlnGluProValLysGlyProValSerThrLysProGlySerCysProIleLeu 20
Db 1554 GCGCAAGAGCCAGTCAAAGGTCCAGTCTCCACTAAGCCTGGTCTCTGCCCATTTCTTG 1613

Qy 21 IleArgCysAlaMetLeuAsnProProAsnArgCysLeuLysAspThrAspCysProGly 40
Db 1614 ATCCGGTGGCCCATGTTGAATCCCCCTAACCCGTGCTTGAAGATAGTACTGACTGCCACGGA 1673

Qy 41 IleLysLysCysCysGluGlySerCysGlyMetAlaCysPheValProGln 57
Db 1674 ATCAAGAAGTGCTGTGAAGGCTCTTGGGGATGGCTGTTCGTTCCCCAG 1724

RESULT 8
ABL66488
ID ABL66488 standard; DNA; 2309 BP.
XX
AC ABL66488;
XX
DT 15-MAY-2002 (first entry)
XX
DE Lung cancer related gene sequence SEQ ID NO:4825.
XX
KW Human; cancer; colon; breast; ovary; oesophagus; kidney; thyroid;

KW stomach; lung; prostate; pancreas; carcinoma; antitumour; cancerous;
KW cytostatic; gene therapy; antineoplastic; Wilm's tumour; adenocarcinoma;
XX gene; ds.
OS Homo sapiens.
XX
PN WO200194629-A2.
XX
PD 13-DEC-2001.
XX
PF 30-MAY-2001; 2001WO-US10838.
XX
PR 05-JUN-2000; 2000US-209473P.
PR 05-JUN-2000; 2000US-209531P.
PR 18-SEP-2000; 2000US-233133P.
PR 18-SEP-2000; 2000US-233617P.
PR 20-SEP-2000; 2000US-234009P.
PR 20-SEP-2000; 2000US-234034P.
PR 20-SEP-2000; 2000US-234052P.
PR 22-SEP-2000; 2000US-234509P.
PR 22-SEP-2000; 2000US-234567P.
PR 25-SEP-2000; 2000US-234923P.
PR 25-SEP-2000; 2000US-234924P.
PR 25-SEP-2000; 2000US-235077P.
PR 25-SEP-2000; 2000US-235082P.
PR 25-SEP-2000; 2000US-235134P.
PR 25-SEP-2000; 2000US-235280P.
PR 26-SEP-2000; 2000US-235637P.
PR 26-SEP-2000; 2000US-235638P.
PR 27-SEP-2000; 2000US-235711P.
PR 27-SEP-2000; 2000US-235720P.
PR 27-SEP-2000; 2000US-235840P.
PR 27-SEP-2000; 2000US-235863P.
PR 28-SEP-2000; 2000US-236028P.
PR 28-SEP-2000; 2000US-236032P.
PR 28-SEP-2000; 2000US-236033P.
PR 28-SEP-2000; 2000US-236034P.
PR 28-SEP-2000; 2000US-236109P.
PR 28-SEP-2000; 2000US-236111P.
PR 29-SEP-2000; 2000US-236842P.
PR 29-SEP-2000; 2000US-236891P.
PR 02-OCT-2000; 2000US-237172P.
PR 02-OCT-2000; 2000US-237173P.
PR 02-OCT-2000; 2000US-237278P.
PR 02-OCT-2000; 2000US-237294P.
PR 02-OCT-2000; 2000US-237295P.
PR 02-OCT-2000; 2000US-237316P.
PR 03-OCT-2000; 2000US-237425P.
PR 03-OCT-2000; 2000US-237598P.
PR 03-OCT-2000; 2000US-237604P.
PR 03-OCT-2000; 2000US-237606P.
PR 03-OCT-2000; 2000US-237608P.
PR 01-NOV-2000; 2000US-244867P.
PR 01-NOV-2000; 2000US-245084P.
XX
XX (AVAL-) AVALON PHARM.
PI Young PE, Augustus M, Carter KC, Ebner R, Endress G, Horrigan S;
PI Soppet DR, Weaver Z;
XX
DR WPI; 2002-188264/24.
XX
XX Screening for anti-neoplastic agent involves exposing cells to a
PT chemical agent to be tested for anti-neoplastic activity, and
PT determining a change in expression of a gene of a signature gene set -
XX
PS Claim 1; SEQ ID 4825; 44pp; English.
XX
CC The present invention describes a method (M1) for screening for an
CC anti-neoplastic agent. The method involves exposing cells to a chemical
CC agent to be tested for anti-neoplastic activity, determining a change in
CC expression of at least one gene (I) of a signature gene set, where (I)
CC comprises a sequence (S) selected from 8447 sequences (given in ABL61664

CC to ABL70110), or is at least 95% identical to (S), where a change in
CC expression is indicative of anti-neoplastic activity. (I) has cytostatic
CC activity and can be used in gene therapy. M1 can be used for screening
CC an anti-neoplastic agent, and can be used for producing a product which
CC is the data collected with respect to the anti-neoplastic agent as a
CC result of M1, and the data is sufficient to convey the chemical
CC structure and/or properties of the agent. M1 can be used in the
CC treatment of cancer such as colon, breast, stomach, lung, thyroid,
CC oesophageal, ovarian, kidney, prostate or pancreatic cancer,
CC adenocarcinoma, carcinoma, clear cell cancer, infiltrating ductal cancer,
CC infiltrating lobular cancer, squamous cell carcinoma, neuroendocrine
CC carcinoma, papillary carcinoma and Wilms' tumour.

SQ Sequence 2309 BP; 595 A; 537 C; 603 G; 572 T; 2 other;

Alignment Scores:
Pred. No.: 4.77e-28 Length: 2309
Score: 327.00 Matches: 57
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 24 Gaps: 0

09-833799-13B (1-57) x ABL66488 (1-2309)

QY 1 AlaGlnGluProValLysGlyProValSerThrLysProGlySerCysProIleLeu 20
Db 1554 GCGAAGAGCCAGTCAAAGGTCAGTCTCCACTAAGCCTGGCTCTGCCCATTTATCTTG 1613

QY 21 IleArgCysAlaMetLeuAsnProProAsnArgCysLeuLysAspThrAspCysProGly 40
Db 1614 ATCCGGTGCAGCCATGTTGATCCCTAACCCTGCTTGAAGATACTGACTGCCAGGA 1673

QY 41 IleLysLysCysCysGluGlySerCysGlyMetAlaCysPheValProGln 57
Db 1674 ATCAAGAAGTGTGTGAAGGCTCTTGGGGATGGCCTGTTTCGTTCCCCAG 1724

RESULT 9
AAQ56674
ID AAQ56674 standard; DNA; 177 BP.

XX AC AAQ56674;
XX 23-SEP-1994 (first entry)
XX Elafin derivative Leu25 DNA.
XX Elafin; derivative; mutation; transformation; E. coli; yeast;
KW Bacillus subtilis; elastase inhibitor; oxidation; ss.

XX Synthetic.
XX WO9404697-A.
XX 03-MAR-1994.
XX 11-AUG-1993; 93WO-JP01133.
XX 11-AUG-1992; 92JP-0234085.
XX (TSUR) TSUMURA & CO.

PI Amagaya S, Ishima Y, Kaji A, Okawa N, Yoshida M;
XX WPI; 1994-083211/10.
DR P-PSDB; AAR48552.

XX Recombinant modified elafin with improved oxidation stability -
PT has pharmaceutical use as an elastase inhibitor
XX Claim 1; Page 23; 35pp; Japanese.

CC The sequences given in AAQ56674-76 encode elafin derivatives. The DNA

CC encoding wild type elafin is mutated, inserted into a suitable
CC vector and then used to transform E. coli, yeast, Bacillus subtilis
CC or animal cells. The modified elafin is expressed when the
CC transformed cells are cultured. The modified elafin are drugs with
CC elastase inhibitor activity. They have improved oxidation stability
CC over natural elafin and thus retain activity better under oxidative
CC conditions.

SQ Sequence 177 BP; 39 A; 49 C; 48 G; 41 T; 0 other;

Alignment Scores:
Pred. No.: 4.55e-29 Length: 177
Score: 324.00 Matches: 56
Percent Similarity: 100.00% Conservative: 1
Best Local Similarity: 98.25% Mismatches: 0
Query Match: 99.08% Indels: 0
DB: 15 Gaps: 0

09-833799-13B (1-57) x AAQ56674 (1-177)

QY 1 AlaGlnGluProValLysGlyProValSerThrLysProGlySerCysProIleLeu 20
Db 1 GCACAGGAACCCAGTTAAAGTCCGGTGTGACCAAAACCGGCTCTTGGCCGATTATCCTG 60

QY 21 IleArgCysAlaMetLeuAsnProProAsnArgCysLeuLysAspThrAspCysProGly 40
Db 61 ATCCGCTGCGCTTGTGTAACCCGCGAACCGTTGTCTGAAAGACACTGACTGCCCGGT 120

QY 41 IleLysLysCysCysGluGlySerCysGlyMetAlaCysPheValProGln 57
Db 121 ATCAAAAATGCTGCGAAGTTCTTTCGGGTATGGCATGCTTCGTTCCGCAG 171

RESULT 10
AAV28852
ID AAV28852 standard; cDNA to mRNA; 8598 BP.

XX AC AAV28852;
XX 04-AUG-1998 (first entry)
XX pPIC9/ELF25L cDNA construct containing the elafin 25L gene.
XX Elafin; elafin 25L gene; Pichia pastoris; expression vector;
KW alcohol oxidase 1; controlling region; promoter; ds.

XX Synthetic.
XX Pichia pastoris.
FH Key Location/Qualifiers
FT CDS 949..1377
FT /*tag= a
FT sig_peptide 949..1203
FT /*tag= b
FT mat_peptide 1204..1374
FT /*tag= c

XX JP10127292-A.
XX 19-MAY-1998.
XX 31-OCT-1996; 96JP-0304233.
XX 31-OCT-1996; 96JP-0304233.

XX (TSUR) TSUMURA & CO.
XX WPI; 1998-340667/30.
DR P-PSDB; AAW57237.

XX New elafin expression vector - used to transform microbial host(s)
PT for production of elafin commercially
XX Example 1; Page 11-15; 21pp; Japanese.

XX The present sequence represents a pPIC9/ELF25L cDNA construct containing
CC the elafin 25L gene, used in an example of the present invention. The
CC present invention describes: (1) an elafin (EL) expression vector
CC comprising a gene encoding EL and a gene encoding alcohol oxidase 1
CC controlling region (AOER) which controls the expression of EL encoding
CC gene; (2) a microbial host transformed with the vector; (3) DNA fragment
CC for recombinant transformation comprising: (a) a first region homologous
CC to the genomic DNA of the host; (b) a promoter region of (AOER) gene; (c)
CC a gene encoding signal peptide; (d) a linker designed to secrete a
CC matured elafin; (e) a gene encoding EL; (f) selection marker gene; (g)
CC stop codon, and (h) a second region homologous to a part of the genomic
CC DNA of the host, and (4) a microbial host transformed with the DNA
CC fragment of (3). The microbial hosts can be used to prepare EL
CC commercially. The production of EL by the microbial hosts can be carried
CC on a large scale.

SQ Sequence 8598 BP; 2306 A; 2025 C; 1920 G; 2347 T; 0 other;

Alignment Scores:

Pred. No.: 5.31e-27 Length: 8598
Score: 324.00 Matches: 56
Percent Similarity: 100.00% Conservative: 1
Best Local Similarity: 98.25% Mismatches: 0
Query Match: 99.08% Indels: 0
DB: 19 Gaps: 0

09-833799-13B (1-57) x AAV28852 (1-8598)

QY 1 AlaGlnGluProValLysGlyProValSerThrLysProGlySerCysProIleIleLeu 20
Db 1204 GCTCAAGAACACAGTTAAGGTCGGTTCGACCAACCGGGCTCTTGCCCGATTATCTG 1263
QY 21 IleArgCysAlaMetLeuAsnProProAsnArgCysLeuLysAspThrAspCysProGly 40
Db 1264 ATCCGCTGCGCTTGCTGAACCCGCCGACCGTGTCTGAAAGACACTGACTGCCCGGGT 1323
QY 41 IleLysLysCysCysGluGlySerCysGlyMetAlaCysPheValProGln 57
Db 1324 ATCAAAAAATGCTGCGAAGGTTCTTTCGGGTATGGCATGCTTCGTTCCGCAG 1374

RESULT 11

AAQ56675
ID AAQ56675 standard; DNA; 177 BP.

XX AC AAQ56675;

XX DT 23-SEP-1994 (first entry)

XX DE Elafin derivative Val25 DNA.

XX KW Elafin; derivative; mutation; transformation; E. coli; yeast;
KW Bacillus subtilis; elastase inhibitor; oxidation; ss.

XX OS Synthetic.

XX PN WO9404697-A.

XX PD 03-MAR-1994.

XX PF 11-AUG-1993; 93WO-JP01133.

XX PR 11-AUG-1992; 92JP-0234085.

XX PA (TSUR) TSUMURA & CO.

XX PI Amagaya S, Ishima Y, Kaji A, Okawa N, Yoshida M;

XX DR WPI; 1994-083211/10.

XX DR P-PSDB; AAR48553.

XX PT Recombinant modified elafin with improved oxidation stability -
has pharmaceutical use as an elastase inhibitor

XX Claim 1; Page 24; 35pp; Japanese.
XX The sequences given in AAQ56674-76 encode elafin derivatives. The DNA
CC encoding wild type elafin is mutated, inserted into a suitable
CC vector and then used to transform E. coli, yeast, Bacillus subtilis
CC or animal cells. The modified elafin is expressed when the
CC transformed cells are cultured. The modified elafin are drugs with
CC elastase inhibitor activity. They have improved oxidation stability
CC over natural elafin and thus retain activity better under oxidative
CC conditions.
XX SQ Sequence 177 BP; 39 A; 50 C; 48 G; 40 T; 0 other;

Alignment Scores:

Pred. No.: 5.94e-29 Length: 177
Score: 323.00 Matches: 56
Percent Similarity: 100.00% Conservative: 1
Best Local Similarity: 98.25% Mismatches: 0
Query Match: 98.78% Indels: 0
DB: 15 Gaps: 0

09-833799-13B (1-57) x AAQ56675 (1-177)

QY 1 AlaGlnGluProValLysGlyProValSerThrLysProGlySerCysProIleIleLeu 20
Db 1 GCACAGGAACACAGTTAAAGGTCGGTTCGACCAACCGGGCTCTTGCCCGATTATCCTG 60
QY 21 IleArgCysAlaMetLeuAsnProProAsnArgCysLeuLysAspThrAspCysProGly 40
Db 61 ATCCGCTGCGCTGTCTGAAACCCGCCGACCGTGTCTGAAAGACACTGACTGCCCGGGT 120
QY 41 IleLysLysCysCysGluGlySerCysGlyMetAlaCysPheValProGln 57
Db 121 ATCAAAAAATGCTGCGAAGGTTCTTTCGGGTATGGCATGCTTCGTTCCGCAG 171

RESULT 12

AAQ56676
ID AAQ56676 standard; DNA; 177 BP.

XX AC AAQ56676;

XX DT 23-SEP-1994 (first entry)

XX DE Elafin derivative Ile25 DNA.

XX KW Elafin; derivative; mutation; transformation; E. coli; yeast;
KW Bacillus subtilis; elastase inhibitor; oxidation; ss.

XX OS Synthetic.

XX PN WO9404697-A.

XX PD 03-MAR-1994.

XX PF 11-AUG-1993; 93WO-JP01133.

XX PR 11-AUG-1992; 92JP-0234085.

XX PA (TSUR) TSUMURA & CO.

XX PI Amagaya S, Ishima Y, Kaji A, Okawa N, Yoshida M;

XX DR WPI; 1994-083211/10.

XX DR P-PSDB; AAR48554.

XX PT Recombinant modified elafin with improved oxidation stability -
has pharmaceutical use as an elastase inhibitor

XX PS Claim 1; Page 25; 35pp; Japanese.

XX The sequences given in AAQ56674-76 encode elafin derivatives. The DNA
CC encoding wild type elafin is mutated, inserted into a suitable

CC vector and then used to transform E. coli, yeast, Bacillus subtilis
CC or animal cells. The modified elafin is expressed when the
CC transformed cells are cultured. The modified elafin are drugs with
CC elastase inhibitor activity. They have improved oxidation stability
CC over natural elafin and thus retain activity better under oxidative
CC conditions.
XX
SQ Sequence 177 BP; 40 A; 49 C; 47 G; 41 T; 0 other;

Alignment Scores:
Pred. No.: 5.94e-29 Length: 177
Score: 323.00 Matches: 56
Percent Similarity: 100.00% Conservative: 1
Best Local Similarity: 98.25% Mismatches: 0
Query Match: 98.78% Indels: 0
DB: 15 Gaps: 0

09-833799-13B (1-57) x AAQ56676 (1-177)

QY 1 AlaGlnGluProValLysGlyProValSerThrLysProGlySerCysProIleLeu 20
Db 1 GCACAGGAACCACTAAAGGTCCGGTGTGACCAACCGGGCTCTGCCCGATTATCTTG 60
QY 21 IleArgCysAlaMetLeuAsnProProAsnArgCysLeuLysAspThrAspCysProGly 40
Db 61 ATCCGCTGCGTATTCTGAACCGCGCAACCGTTGCTGAAGACACTGACTGCCCGGT 120
QY 41 IleLysLysCysCysGluGlySerCysGlyMetAlaCysPheValProGln 57
Db 121 ATCAAAAATGCTGCAAGTTCTTGGGTATGGCATGCTTCGTTCCGCAG 171

RESULT 13

AAQ28757
ID AAQ28757 standard; DNA; 321 BP.

XX
AC AAQ28757;

DT 25-FEB-1993 (first entry)

DE Partial sequence of tumour suppressor gene U9.

XX Can19; tumour suppressor gene; cancer; therapy; ss.

XX Homo sapiens.

XX WO9215602-A.

XX 17-SEP-1992.

XX 28-FEB-1992; 92WO-US01624.

XX 28-FEB-1991; 91US-0662216.

XX (DAND) DANA FABER CANCER INST INC.

XX Sager R;

XX WPI; 1992-331663/40.

XX Diagnoses and treatment of cancer - using candidate tumor suppressor
XX genes or the corresp. antibodies.

XX Claim 29; Page 37-38; 54pp; English.

XX An adaptation of the subtractive hybridization technique was used
XX which utilizes a biotinylation-based subtraction procedure instead
XX of hydroxyapatite as previously used. In this procedure, a single
XX strand phagemid cDNA library from normal cell polyA+ mRNA is
XX hybridized with excess biotinylated tumor polyA+ mRNA, and the
XX resulting double stranded sequences are removed by binding to
XX streptavidin. The remaining single-stranded phagemid cDNAs are
XX converted to double-stranded form and used to transform bacterial
XX host cells. The resulting subtracted cDNA library is differentially

CC screened with total cDNA from normal and tumor cells. This method
CC produced some 20 additional cloned cDNAs. Also found by this
CC method were several genes which, on the basis of the partial DNA
CC sequences appear to be novel sequences not previously entered
CC into GENBANK. The portion of the cDNAs so sequenced represents
CC part of the coding region and/or part of the 3' untranslated region
CC of each cDNA (see Q28749-58).
XX

SQ Sequence 321 BP; 81 A; 80 C; 84 G; 76 T; 0 other;

Alignment Scores:
Pred. No.: 1.61e-28 Length: 321
Score: 322.00 Matches: 56
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 98.47% Indels: 0
DB: 13 Gaps: 0

09-833799-13B (1-57) x AAQ28757 (1-321)

QY 1 AlaGlnGluProValLysGlyProValSerThrLysProGlySerCysProIleLeu 20
Db 144 GCGCAAGAGCCAGTCAAGGTCCAGTCTCCACTAAGCCTGGCTCCTGCCCATTTCTTG 203

QY 21 IleArgCysAlaMetLeuAsnProProAsnArgCysLeuLysAspThrAspCysProGly 40
Db 204 ATCCGCTGCGCATGTTGAATCCTCCTAACCGCTGCTTGAAGATACTGACTGCCAGGA 263

QY 41 IleLysLysCysCysGluGlySerCysGlyMetAlaCysPheValPro 56

Db 264 ATCAAGAAGTGTGTGAAGGCTCTTGGGGATGGCCTGTTTCGTTCCC 311

RESULT 14

AAQ44862

ID AAQ44862 standard; cDNA to mRNA; 737 BP.

XX
AC AAQ44862;

DT 17-OCT-1994 (first entry)

XX SPAl gene.

XX SPAl; sodium ion, potassium ion, ATP-ase inhibitor peptide; cardiac;
XX vascular disease; PCR; polymerase chain reaction; ss.

XX Sus scrofa.

XX Key Location/Qualifiers
XX CDS 1..570

XX /*tag= a

XX /product= SPAl

XX JP06049098-A.

XX 22-FEB-1994.

XX 29-JUL-1992; 92JP-0202286.

XX 29-JUL-1992; 92JP-0202286.

XX (EISA) EISAI CO LTD.

XX WPI; 1994-097819/12.

XX P-PSDB; AAR50334.

XX Sodium ion, potassium ion ATP-ase inhibitor peptide and corresp.
XX DNA - useful to treat cardiac and vascular disease

XX Claim 1; Page 6-7; 8pp; Japanese.

XX The sequence (AAQ44862) encodes the SPAl peptide which is useful
XX as a drug for treating cardiac and vascular diseases. The peptide
XX is found as a pre-sequence (AAR50335) which is amplified using

cancer metastasis associated protein; ss.

Mus musculus.

Key Location/Qualifiers
CDS 40..420
/*tag= a

WO9845431-A1.
15-OCT-1998.
07-APR-1998; 98WO-JP01592.
08-APR-1997; 97JP-0105333.
(BANY) BANYU PHARM CO LTD.
Arakawa H, Morita M, Ohta M;
WPI; 1999-080732/07.
P-PSDB; AAW98909.

Protein associated with cancer metastasis and gene encoding it -
useful for screening for potential inhibitors of cancer metastasis

Claim 2; Page 46-47; 74pp; Japanese.

The present invention provides gene sequences associated with cancer
metastasis which are isolated from mouse IMC carcinoma cells by
detection of their higher expression in IMC-HM cell lines than in
IMC-LM cell lines using differential display of the mRNA in these cells.
The gene sequences can be used for the screening of potential inhibitors
of cancer metastasis by either: bringing into contact with the cancer
metastasis associated protein (CMAP) and determining the degree of
binding; or creating a transformant cell line which expresses CMAP and
measuring the degree of expression of CMAP using an antibody recognising
the protein, either in the presence or absence of the potential
inhibitor. IMC-HM cells transformed with antisense CMAP DNA show a
lowered ability to metastasise. The present sequence represents a
specifically claimed gene sequence from the present invention.

Sequence 684 BP; 175 A; 176 C; 179 G; 154 T; 0 other;

Alignment Scores:
Pred. No.: 1.57e-09 Length: 684
Score: 161.00 Matches: 28
Percent Similarity: 58.49% Conservative: 3
Best Local Similarity: 52.83% Mismatches: 22
Query Match: 49.24% Indels: 0
DB: 20 Gaps: 0

09-833799-13B (1-57) x AAX18516 (1-684)

Qy 4 ProValLysGlyProValSerThrLysProGlySerCysProIleLeuIleArgCys 23
|||::: ||||| ||||||| ||| |||||
Db 256 CCCATTCCGAACAGTGTGGAGGAGCGCTGGAGGTGCGTCAAAACTCAGGCAAGATGT 315
|||::: ||||| ||||||| ||| |||||

Qy 24 AlaMetLeuAsnProAsnArgCysLeuLysAspThrAspCysProGlyIleLysLys 43
||||| ||||||| ||||| ||| ::||| ||| |||
Db 316 ATGATGCTTAACCTCCCAATGTCTGCCAGAGGACGGGCAGTGTCACGGCAAATACAAG 375
||||| ||||||| ||||||| ||| |||||

Qy 44 CysCysGluGlySerCysGlyMetAlaCysPheValPro 56
||||| ||||||| ||||||| ||| |||||
Db 376 TGCTGTGAGGGTATATGTGGGAAAGTCTGCCTGCCCCCG 414
||||| ||||||| ||||||| ||| |||||

RESULT 18
AAX18514
ID AAX18514 standard; cDNA to mRNA; 691 BP.
XX AC
XX AAX18514;
DT 05-MAY-1999 (first entry)


```

AC AAC97526;
XX
DT 27-FEB-2001 (first entry)
XX
DE DNA encoding a serine protease inhibitory peptide.
XX
KW Serine protease inhibitor; cytostatic; anti-inflammatory; arthritis;
KW protease mediated tissue destruction; emphysema; glomerulonephritis;
KW periodontitis; muscular dystrophy; tumour invasion; chymotrypsin;
KW elastase; ds.
XX
OS Synthetic.
XX
PN US6132990-A.
XX
PD 17-OCT-2000.
XX
PF 07-JUN-1991; 91US-0712354.
XX
PR 03-JAN-1989; 89US-0293042.
PR 06-DEC-1984; 84US-0678822.
PR 02-DEC-1985; 85US-0803471.
PR 29-JUL-1986; 86US-0890526.
PR 30-MAR-1987; 87US-0031846.
PR 04-AUG-1987; 87US-0082962.
XX
PA (AMGE-) AMGEN BOULDER INC.
XX
PI Bandyopadhyay PK, Eisenberg SP, Stetler GL, Thompson RC;
XX WPI; 2000-678667/66.
XX
PT New serine protease inhibitors and DNA sequences for treating a
PT protease-mediated condition or tissue destruction e.g. emphysema or
PT tumor invasion and for recombinant production of inhibitors -
XX
PS Claim 22; Column 61-62; 47pp; English.
XX
CC This invention relates to new purified and isolated mammalian serine
CC protease inhibitor proteins which comprise at least 8 cysteine residues
CC and no more than 107 amino acids. The protease inhibitors are capable of
CC inhibiting chymotrypsin and elastase. Sequences AAC97526 - AAC97534,
CC AAC97574 - AAC97581 and AAC97614 - AAC97641 represent DNA encoding the
CC protease inhibitors of the invention, and include oligonucleotide
CC sequences used in the isolation and characterisation of the proteins.
CC Primers AAC97535 - AAC97573 and AAC95582 - AAC97613 are used in the
CC construction of DNA encoding the protease inhibitors. Peptide sequences
CC AAB53098 - AAB53122 represent the protease inhibitor of the invention and
CC various other peptides used in the isolation of the protease inhibitors.
CC The protease inhibitors have cytosolic and anti-inflammatory activity.
CC The serine protease inhibitor protein is useful for treating a
CC protease-mediated condition, which includes protease mediated tissue
CC destruction, e.g. emphysema, arthritis, glomerulonephritis,
CC periodontitis, muscular dystrophy or tumour invasion. It is also useful
CC for modulating protease activity. The DNAs are useful for producing the
CC proteins, especially by recombinant methods.
XX
SQ Sequence 180 BP; 42 A; 44 C; 50 G; 44 T; 0 other;

Alignment Scores:
Pred. No.: 1.65e-08 Length: 180
Score: 146.00 Matches: 27
Percent Similarity: 55.36% Conservative: 4
Best Local Similarity: 48.21% Mismatches: 23
Query Match: 44.65% Indels: 2
DB: 21 Gaps: 1

09-833799-13B (1-57) x AAC97526 (1-180)

QY 3 GluProValLysGlyProValSerThr-----LysProGlySerCysProIleLeu 20
Db 4 GATCCTGTTCACCCCCAACACCACAGGAGGCCTGGGAAGTGCCAGTGACTTAT 63
:::||||| ||| ||||| |

```


XX Sugiyama T, Kamimura T, Masuda K, Okada M, Ohtsuka E;
PI WPI; 1989-220549/30.
XX P-PSDB; AAP90126, AAP90384.
DR
XX Elastase-inhibiting peptide of low serine protease-inhibiting activity
PT - contains C-terminal portion of human polymorphonuclear leukocyte
PT elastase inhibiting protein.
XX
PS Disclosure; fig 2; 55pp; Japanese.
XX
CC The sequence encodes a fragment of human polymorphonuclear leukocyte
CC elastase inhibiting protein (Asn 55-Ala 107). See AAP90126 for encoded
CC peptide. Misc. features a-g are BamHI, Sall, MluI, NdeI, BglII, XhoI and
CC PstI sites resp. See also AAP90384.
XX
SQ Sequence 194 BP; 48 A; 42 C; 53 G; 51 T; 0 other;
XX
Alignment Scores:
Pred. No.: 4.01e-08 Length: 194
Score: 143.00 Matches: 27
Percent Similarity: 55.36% Conservative: 4
Best Local Similarity: 48.21% Mismatches: 23
Query Match: 43.73% Indels: 2
DB: 10 Gaps: 1
09-833799-13B (1-57) x AAN90354 (1-194)
QY 3 GluProValLysGlyProValSerThr-----LysProGlySerCysProIleLeu 20
Db 1 GATCGGTGACACCCGACCGGACGCGTGTAAACCGGTAATGTCCGGTACATAT 60
QY 21 IleArgCysAlaMetLeuAsnProAsnArgCysLeuLysAspThrAspCysProGly 40
Db 61 GGTCAGTGCTGTGCTGAACCGCGCAACTTCTGTGAATGGACGGTCAGGTAAACGA 120
QY 41 IleLysLysCysCysGluGlySerCysGlyMetAlaCysPheValPro 56
Db 121 GATCTGAATGTTGTATGGGTATGTGGTAAATCTTGTGTTCTCCG 168
RESULT 24
ID ABK88016 standard; DNA; 321 BP.
AC ABK88016;
XX 07-OCT-2002 (first entry)
XX DNA encoding human secretory leukoprotease inhibitor (SLPI) protein.
KW Secretory leukoprotease inhibitor; human; gene; ds; protease inhibitor;
KW malaria; emphysema; asthma; chronic obstructive pulmonary disease; SLPI;
KW cystic fibrosis; otitis media; otitis externa; HIV; psoriasis; eczema;
KW human immunodeficiency virus; atopic dermatitis; muscular dystrophy;
KW herpes; ulceration; sepsis; rheumatoid arthritis; periodontal disease;
KW tumour metastasis; tumour angiogenesis; osteoporosis; Paget's disease;
KW glomerulonephritis; scleroderma; Alzheimer's disease; hypertension.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT CDS 1..321
FT /*tag= a
FT /product= "Secretory leukoprotease inhibitor"
FT /partial
FT /note= "No start or stop codon shown"
XX
PN WO200250287-A2.
XX
PD 27-JUN-2002.
XX
PF 18-DEC-2001; 2001WO-US49256.

XX 18-DEC-2000; 2000US-256699P.
PR 20-NOV-2001; 2001US-331966P.
XX
PA (ARRI-) ARRIVA PHARM INC.
XX
PI Barr PJ, Gibson HL, Pemberton P;
XX
DR WPI; 2002-500631/53.
DR P-PSDB; AAU99874.
XX
PT Novel fusion protein useful for inhibiting protease activity associated
PT with a disorder such as emphysema, asthma, comprises a first protease
PT inhibitor comprising alpha 1-antitrypsin and a second protease
PT inhibitor -
XX
PS Disclosure; Page 32; 134pp; English.
XX
CC This invention relates to a novel fusion protein comprising a first
CC protease inhibitor comprising an alpha1-antitrypsin or its functionally
CC active portion and a second protease inhibitor or its functionally
CC active protein. The fusion proteins of the invention may act as an
CC inhibitor of protease activity. The fusion protein of the invention
CC is useful for inhibiting protease activity associated with a disorder
CC such as emphysema, asthma, chronic obstructive pulmonary disease,
CC cystic fibrosis, otitis media, otitis externa or HIV infection, or
CC for treating an individual suffering from or at risk for a disease or
CC disorder involving unwanted protease activity. The proteins are useful
CC for treating dermatological diseases such as atopic dermatitis, eczema
CC and psoriasis, in inflammatory responses to viral infection, and for
CC treating herpes infection, corneal or epidermal ulceration, chronic
CC non-healing wounds, sepsis, rheumatoid arthritis, periodontal disease,
CC tumour metastasis and tumour angiogenesis, gastric ulceration,
CC osteoporosis, Paget's disease, glomerulonephritis, scleroderma, malaria,
CC bacterial infection, Alzheimer's disease, hypertension and muscular
CC dystrophy. The present sequence represents the DNA encoding the human
CC secretory leukoprotease inhibitor used to create the fusion protein
CC of the invention.
XX
SQ Sequence 321 BP; 92 A; 67 C; 80 G; 82 T; 0 other;
XX
Alignment Scores:
Pred. No.: 7.43e-08 Length: 321
Score: 143.00 Matches: 27
Percent Similarity: 55.36% Conservative: 4
Best Local Similarity: 48.21% Mismatches: 23
Query Match: 43.73% Indels: 2
DB: 24 Gaps: 1
09-833799-13B (1-57) x ABK88016 (1-321)
QY 3 GluProValLysGlyProValSerThr-----LysProGlySerCysProIleLeu 20
Db 145 GACCCAGTTGACACCCCAACCACTAGAGAAGCCAGTAAAGTGTCCAGTTACTTAC 204
QY 21 IleArgCysAlaMetLeuAsnProAsnArgCysLeuLysAspThrAspCysProGly 40
Db 205 GGTCATGTTGTATGTTGAACCCCACTTCTGTGAATGGACGGTCAATGTAAGAGA 264
QY 41 IleLysLysCysCysGluGlySerCysGlyMetAlaCysPheValPro 56
Db 265 GACTTGAAGTGTGTATGGGTATGTGTGGTAAAGTCTGTGTTTCCCA 312
RESULT 25
AAN60463
ID AAN60463 standard; DNA; 324 BP.
XX
AC AAN60463;
XX
DT 01-JAN-1980 (first entry)
XX
DE Synthetic sequence capable of directing microbial synthesis of a
DE serine protease inhibitor having similar properties to protein

DE isolated from parotid secretions.
XX Serine protease-inhibitor; parotid; leukocyte-elastase-inhibitor;
KW trypsin-; inhibitor; ss.
XX
OS Synthetic.
XX
PN WO8603519-A.
XX
PD 19-JUN-1986.
XX
PF 04-DEC-1985; 85WO-US02385.
XX
PR 04-DEC-1985; 85WO-US02385.
PR 06-DEC-1984; 84US-0678822.
PR 02-DEC-1985; 85US-0803471.
XX
PA (SYNE-) SYNERGEN BIOLOG INC.
XX
PI Bandyopadh PK, Eisenberg SP, Stetler GL, Thompson RC;
XX
DR WPI; 1986-169458/26.
DR P-PSDB; AAP60562.
XX
PT New synthetic DNA sequences for directing microbial synthesis -
PT for prodn. of single poly:peptide chain serine protease inhibitor
PT having leukocyte elastase and trypsin inhibitory sites
XX
PS Disclosure; Page 14; 59pp; English.
XX
CC The sequence directs synthesis of a single chain polypeptide serine
CC protease-inhibitor, which believed to have at least 2 active sites,
CC 1 exhibiting leukocyte-elastase-inhibiting properties and the other
CC exhibiting activity against trypsin. See also AAN60464-69 and AAP60562-
CC 66.
XX
SQ Sequence 324 BP; 87 A; 85 C; 88 G; 64 T; 0 other;
XX
Alignment Scores:
Pred. No.: 7.52e-08 Length: 324
Score: 143.00 Matches: 27
Percent Similarity: 55.36% Conservative: 4
Best Local Similarity: 48.21% Mismatches: 23
Query Match: 43.73% Indels: 2
DB: 7 Gaps: 1
09-833799-13B (1-57) x AAN60463 (1-324)
QY 3 GluProValLysGlyProValSerThr-----LysProGlySerCysProIleIleLeu 20
Db 145 GATCCGGTTGATACCCCGAACCCGACTCGTCGAAACCGGGTAAATGCCCGGTAACTAT 204
QY 21 IleArgCysAlaMetLeuAsnProProAsnArgCysLeuLysAspThrAspCysProGly 40
Db 205 GGCCAGTGTCTGATGCTGAACCCCGCAACTTCTGCGAAATGGACGGCCAGTGTAAACGA 264
QY 41 IleLysLysCysCysGluGlySerCysGlyMetAlaCysPheValPro 56
Db 265 GATCTGAAATGCTGTATGGGTATGTGCGGCAAAATCTTGTGTTTCCCG 312
RESULT 26
AAN60464
ID AAN60464 standard; DNA; 324 BP.
XX
AC AAN60464;
XX
DT 01-JAN-1980 (first entry)
XX
DE Synthetic sequence capable of directing microbial synthesis of a
DE secretory leukocyte protease-inhibitor.
XX
KW Serine protease-inhibitor; parotid; leukocyte-elastase-inhibitor;
KW trypsin-; inhibitor; ss.

XX
OS Synthetic.
XX
PN WO8603519-A.
XX
PD 19-JUN-1986.
XX
PF 04-DEC-1985; 85WO-US02385.
XX
PR 04-DEC-1985; 85WO-US02385.
PR 06-DEC-1984; 84US-0678822.
PR 02-DEC-1985; 85US-0803471.
XX
PA (SYNE-) SYNERGEN BIOLOG INC.
XX
PI Bandyopadh PK, Eisenberg SP, Stetler GL, Thompson RC;
XX
DR WPI; 1986-169458/26.
DR P-PSDB; AAP60563.
XX
PT New synthetic DNA sequences for directing microbial synthesis -
PT for prodn. of single poly:peptide chain serine protease inhibitor
PT having leukocyte elastase and trypsin inhibitory sites
XX
PS Disclosure; Page 15; 59pp; English.
XX
CC The sequence directs synthesis of a secretory leukocyte protease-
CC inhibitor. See also AAN60463, AAN60465-69 and AAP60562. AAP60564-66.
XX
SQ Sequence 324 BP; 86 A; 85 C; 87 G; 66 T; 0 other;
XX
Alignment Scores:
Pred. No.: 7.52e-08 Length: 324
Score: 143.00 Matches: 27
Percent Similarity: 55.36% Conservative: 4
Best Local Similarity: 48.21% Mismatches: 23
Query Match: 43.73% Indels: 2
DB: 7 Gaps: 1
09-833799-13B (1-57) x AAN60464 (1-324)
QY 3 GluProValLysGlyProValSerThr-----LysProGlySerCysProIleIleLeu 20
Db 145 GATCCGGTTGATACCCCGAACCCGACTCGTCGAAACCGGGTAAATGCCCGGTAACTAT 204
QY 21 IleArgCysAlaMetLeuAsnProProAsnArgCysLeuLysAspThrAspCysProGly 40
Db 205 GGCCAGTGTCTGATGCTGAACCCCGCAACTTCTGCGAAATGGACGGCCAGTGTAAACGA 264
QY 41 IleLysLysCysCysGluGlySerCysGlyMetAlaCysPheValPro 56
Db 265 GATCTGAAATGCTGTATGGGTATGTGCGGCAAAATCTTGTGTTTCCCG 312
RESULT 27
AAX16194
ID AAX16194 standard; DNA; 324 BP.
XX
AC AAX16194;
XX
DT 19-APR-1999 (first entry)
XX
DE Serine protease inhibitor direct manufacturing DNA sequence.
XX
KW Serine protease inhibitor; human; emphysema; arthritis; periodontitis;
KW muscular dystrophy; tumour invasion; glomerulonephritis; sepsis;
KW acute leukemia; ss.
XX
OS Homo sapiens.
OS Synthetic.
XX
PN US5871956-A.
XX
PD 16-FEB-1999.

```
XX PF 22-JUL-1994; 94US-0279056.
XX PR 30-MAR-1987; 87US-0031846.
XX PR 05-DEC-1984; 84US-0678822.
XX PR 29-JUL-1986; 86US-0890526.
XX PR 03-SEP-1986; 86US-0903471.
XX PR 06-AUG-1990; 90US-0563832.
XX PR 22-JUL-1994; 94US-0279056.
XX PA (AMGE-) AMGEN INC.
XX PI Bandyopadhyay PK, Eisenberg SP, Stetler GL, Thompson RC;
XX DR WPI; 1999-166640/14.
XX PT New DNA sequence encoding mammalian serine protease inhibitor -
XX PT useful for recombinantly producing inhibitors with different
XX PT specificities and treating diseases such as emphysema, arthritis,
XX PT muscular dystrophy, and tumour invasion
XX PS Example 1; Column 23; 37pp; English.
XX CC The present invention describes a DNA sequence (A) encoding an analogue
XX CC of a mammalian serine protease inhibitor (B). The DNA sequences and
XX CC recombinant methods allow manufacture of a class of inhibitors of e.g.
XX CC cathepsin G, elastase, and trypsin, with different specificities. The
XX CC recombinant serine protease product can be directed to act
XX CC intracellularly or extracellularly and is useful in treating conditions
XX CC caused by a disturbance in the native protease/protease inhibitor
XX CC balance such as emphysema, arthritis, glomerulonephritis, peridontitis,
XX CC muscular dystrophy, tumour invasion, sepsis, and acute leukemia. (A)
XX CC allows (B) to be recombinantly produced in sufficient quantities and
XX CC purities so as to provide economical pharmaceutical compositions. (B) is
XX CC resistant to heat, acid, and a variety of proteolytic enzymes, is
XX CC thermodynamically stable in extracellular conditions, and exhibits a
XX CC high degree of self assembly forming an active tertiary structure in the
XX CC absence of biochemical stimuli. The present sequence represents a DNA
XX CC sequence which is used to direct manufacture of recombinant serine
XX CC protease inhibitors.
XX SQ Sequence 324 BP; 87 A; 85 C; 88 G; 64 T; 0 other;

Alignment Scores:
Pred. No.: 7.52e-08 Length: 324
Score: 143.00 Matches: 27
Percent Similarity: 55.36% Conservative: 4
Best Local Similarity: 48.21% Mismatches: 23
Query Match: 43.73% Indels: 2
DB: 20 Gaps: 1

09-833799-13B (1-57) x AAX16194 (1-324)

QY 3 GluProValLysGlyProValSerThr-----LysProGlySerCysProIleIleLeu 20
Db 145 GATCCGGTTGATACCCCGAACCCGACTCGTCGAAACCCGGTAAATGCCCGTAACCTAT 204

QY 21 IleArgCysAlaMetLeuAsnProProAsnArgCysLeuLysAspThrAspCysProGly 40
Db 205 GGCCAGTGTCTGATGCTGAACCCCGCGAACTTCTGCGAAATGGACGCCAGTGTAACGA 264

QY 41 IleLysLysCysGluGlySerCysGlyMetAlaCysPheValPro 56
Db 265 GATCTGAATGCTGTATGGGTATGTGCGCAAAATCTTGTGTTCCCG 312

RESULT 28
AAX16236
ID AAX16236 standard; DNA; 324 BP.
XX AC AAX16236;
XX AC AAX16236;
XX DT 19-APR-1999 (first entry)
XX
```

```
DE XX Serine protease inhibitor direct manufacturing DNA sequence (SLPI).
KW KW Serine protease inhibitor; human; emphysema; arthritis; peridontitis;
KW KW muscular dystrophy; tumour invasion; glomerulonephritis; sepsis;
XX OS acute leukemia; ss.
OS OS Homo sapiens.
OS OS Synthetic.
XX PN US5871956-A.
XX PD 16-FEB-1999.
XX PF 22-JUL-1994; 94US-0279056.
XX PR 30-MAR-1987; 87US-0031846.
XX PR 05-DEC-1984; 84US-0678822.
XX PR 29-JUL-1986; 86US-0890526.
XX PR 03-SEP-1986; 86US-0903471.
XX PR 06-AUG-1990; 90US-0563832.
XX PR 22-JUL-1994; 94US-0279056.
XX PA (AMGE-) AMGEN INC.
XX PI Bandyopadhyay PK, Eisenberg SP, Stetler GL, Thompson RC;
XX DR WPI; 1999-166640/14.
XX PT New DNA sequence encoding mammalian serine protease inhibitor -
XX PT useful for recombinantly producing inhibitors with different
XX PT specificities and treating diseases such as emphysema, arthritis,
XX PT muscular dystrophy, and tumour invasion
XX PS Example 2; Column 29; 37pp; English.
XX CC The present invention describes a DNA sequence (A) encoding an analogue
XX CC of a mammalian serine protease inhibitor (B). The DNA sequences and
XX CC recombinant methods allow manufacture of a class of inhibitors of e.g.
XX CC cathepsin G, elastase, and trypsin, with different specificities. The
XX CC recombinant serine protease product can be directed to act
XX CC intracellularly or extracellularly and is useful in treating conditions
XX CC caused by a disturbance in the native protease/protease inhibitor
XX CC balance such as emphysema, arthritis, glomerulonephritis, peridontitis,
XX CC muscular dystrophy, tumour invasion, sepsis, and acute leukemia. (A)
XX CC allows (B) to be recombinantly produced in sufficient quantities and
XX CC purities so as to provide economical pharmaceutical compositions. (B) is
XX CC resistant to heat, acid, and a variety of proteolytic enzymes, is
XX CC thermodynamically stable in extracellular conditions, and exhibits a
XX CC high degree of self assembly forming an active tertiary structure in the
XX CC absence of biochemical stimuli. The present sequence represents a DNA
XX CC sequence which is used to direct manufacture of recombinant serine
XX CC protease inhibitors.
XX SQ Sequence 324 BP; 86 A; 85 C; 87 G; 66 T; 0 other;

Alignment Scores:
Pred. No.: 7.52e-08 Length: 324
Score: 143.00 Matches: 27
Percent Similarity: 55.36% Conservative: 4
Best Local Similarity: 48.21% Mismatches: 23
Query Match: 43.73% Indels: 2
DB: 20 Gaps: 1

09-833799-13B (1-57) x AAX16236 (1-324)

QY 3 GluProValLysGlyProValSerThr-----LysProGlySerCysProIleIleLeu 20
Db 145 GATCCGGTTGATACCCCGAACCCGACTCGTCGAAACCCGGTAAATGCCCGTAACCTAT 204

QY 21 IleArgCysAlaMetLeuAsnProProAsnArgCysLeuLysAspThrAspCysProGly 40
Db 205 GGCCAGTGTCTGATGCTGAACCCCGCGAACTTCTGCGAAATGGACGCCAGTGTAACGA 264
```

QY 41 IleLysLysCysCysGluGlySerCysGlyMetAlaCysPheValPro 56
AAC97527
Db 265 GATCTGAAATGCTGTAAGGTATGTGCGGCAAAATCTTGTGTTCCCG 312

RESULT 29
AAC97527
ID AAC97527 standard; DNA; 324 BP.
XX
AC AAC97527;
XX
DT 27-FEB-2001 (first entry)
XX
DE DNA encoding a serine protease inhibitor peptide sequence.

XX Serine protease inhibitor; cytostatic; anti-inflammatory; arthritis;
KW protease mediated tissue destruction; emphysema; glomerulonephritis;
KW periodontitis; muscular dystrophy; tumour invasion; chymotrypsin;
KW elastase; ds.
XX

OS Synthetic.
XX
PN US6132990-A.
XX
PD 17-OCT-2000.
XX
PF 07-JUN-1991; 91US-0712354.
XX
PR 03-JAN-1989; 89US-0293042.
PR 06-DEC-1984; 84US-0678822.
PR 02-DEC-1985; 85US-0803471.
PR 29-JUL-1986; 86US-0890526.
PR 30-MAR-1987; 87US-0031846.
PR 04-AUG-1987; 87US-0082962.
XX

XX (AMGE-) AMGEN BOULDER INC.
PA Bandyopadhyay PK, Eisenberg SP, Stetler GL, Thompson RC;
PI WPI; 2000-678667/66.
XX
DR New serine protease inhibitors and DNA sequences for treating a
PT protease-mediated condition or tissue destruction e.g. emphysema or
PT tumor invasion and for recombinant production of inhibitors -
XX
PS Claim 34; Column 64-66; 47pp; English.
XX

CC This invention relates to new purified and isolated mammalian serine
CC protease inhibitor proteins which comprise at least 8 cysteine residues
CC and no more than 107 amino acids. The protease inhibitors are capable of
CC inhibiting chymotrypsin and elastase. Sequences AAC97526 - AAC97534,
CC AAC97574 - AAC97581 and AAC97614 - AAC97641 represent DNA encoding the
CC protease inhibitors of the invention, and include oligonucleotide
CC sequences used in the isolation and characterisation of the proteins.
CC Primers AAC97535 - AAC97573 and AAC95582 - AAC97613 are used in the
CC construction of DNA encoding the protease inhibitors. Peptide sequences
CC AAB53098 - AAB53122 represent the protease inhibitor of the invention and
CC various other peptides used in the isolation of the protease inhibitors.
CC The serine protease inhibitors have cytostatic and anti-inflammatory activity.
CC The serine protease inhibitor protein is useful for treating a
CC protease-mediated condition, which includes protease mediated tissue
CC destruction, e.g. emphysema, arthritis, glomerulonephritis,
CC periodontitis, muscular dystrophy or tumour invasion. It is also useful
CC for modulating protease activity. The DNAs are useful for producing the
CC proteins, especially by recombinant methods.
XX

SQ Sequence 324 BP; 87 A; 85 C; 88 G; 64 T; 0 other;

Alignment Scores:

Pred. No.:	7.52e-08	Length:	324
Score:	143.00	Matches:	27
Percent Similarity:	55.36%	Conservative:	4
Best Local Similarity:	48.21%	Mismatches:	23
Query Match:	43.73%	Indels:	2

DB: 21 Gaps: 1
09-833799-13B (1-57) x AAC97527 (1-324)
QY 3 GluProValLysGlyProValSerThr-----LysProGlySerCysProIleLeu 20
Db 145 GATCCGGTTGATACCCCGAACCCGACTCGTCGAAACCGGGTAAATGCCCGTAACCTAT 204
QY 21 IleArgCysAlaMetLeuAsnProProAsnArgCysLeuLysAspThrAspCysProGly 40
Db 205 GGCCAGTGTCTGATGCTGAACCCCGCAACTTCTGCGAAATGGACGGCAGTGTAAACGA 264
QY 41 IleLysLysCysCysGluGlySerCysGlyMetAlaCysPheValPro 56
Db 265 GATCTGAAATGCTGTATGGGTATGTGCGGCAAAATCTTGTGTTCCCG 312

RESULT 30
AAC97528
ID AAC97528 standard; DNA; 324 BP.
XX
AC AAC97528;
XX
DT 27-FEB-2001 (first entry)
XX
DE Secretory leukocyte protease inhibitor (SLPI) DNA sequence.

XX Serine protease inhibitor; cytostatic; anti-inflammatory; arthritis;
KW protease mediated tissue destruction; emphysema; glomerulonephritis;
KW periodontitis; muscular dystrophy; tumour invasion; chymotrypsin;
KW elastase; ds.
XX

OS Synthetic.

XX US6132990-A.

XX 17-OCT-2000.

XX 07-JUN-1991; 91US-0712354.

XX 03-JAN-1989; 89US-0293042.

XX 06-DEC-1984; 84US-0678822.

XX 02-DEC-1985; 85US-0803471.

XX 29-JUL-1986; 86US-0890526.

XX 30-MAR-1987; 87US-0031846.

XX 04-AUG-1987; 87US-0082962.

XX (AMGE-) AMGEN BOULDER INC.

XX Bandyopadhyay PK, Eisenberg SP, Stetler GL, Thompson RC;
XX WPI; 2000-678667/66.

XX New serine protease inhibitors and DNA sequences for treating a
XX protease-mediated condition or tissue destruction e.g. emphysema or
XX tumor invasion and for recombinant production of inhibitors -

XX Disclosure; Column 5; 47pp; English.

XX This invention relates to new purified and isolated mammalian serine
XX protease inhibitor proteins which comprise at least 8 cysteine residues
XX and no more than 107 amino acids. The protease inhibitors are capable of
XX inhibiting chymotrypsin and elastase. Sequences AAC97526 - AAC97534,
XX AAC97574 - AAC97581 and AAC97614 - AAC97641 represent DNA encoding the
XX protease inhibitors of the invention, and include oligonucleotide
XX sequences used in the isolation and characterisation of the proteins.
XX Primers AAC97535 - AAC97573 and AAC95582 - AAC97613 are used in the
XX construction of DNA encoding the protease inhibitors. Peptide sequences
XX AAB53098 - AAB53122 represent the protease inhibitor of the invention and
XX various other peptides used in the isolation of the protease inhibitors.
XX The serine protease inhibitors have cytostatic and anti-inflammatory activity.
XX The serine protease inhibitor protein is useful for treating a
XX protease-mediated condition, which includes protease mediated tissue
XX destruction, e.g. emphysema, arthritis, glomerulonephritis,

CC periodontitis, muscular dystrophy or tumour invasion. It is also useful
CC for modulating protease activity. The DNAs are useful for producing the
CC proteins, especially by recombinant methods.

XX
SQ Sequence 324 BP; 86 A; 85 C; 87 G; 66 T; 0 other;

Alignment Scores:
Pred. No.: 7.52e-08 Length: 324
Score: 143.00 Matches: 27
Percent Similarity: 55.36% Conservative: 4
Best Local Similarity: 48.21% Mismatches: 23
Query Match: 43.73% Indels: 2
DB: 21 Gaps: 1

09-833799-13B (1-57) x AAC97528 (1-324)

Qy 3 GluProValLysGlyProValSerThr-----LysProGlySerCysProIleLeu 20
AAAC97579
Db 145 GATCCGGTTGATACCCCGAACCCGACTCGTCGAAACCCGGTAAATGCCCGTAACCTAT 204

Qy 21 IleArgCysAlaMetLeuAsnProProAsnArgCysLeuLysAspThrAspCysProGly 40
Db 205 GGCCAGTGTCTGATGCTGAACCCCGCGAACTTCTGCGAAATGGACGCCAGTGTAACCGA 264

Qy 41 IleLysLysCysCysGluGlySerCysGlyMetAlaCysPheValPro 56
Db 265 GATCTGAAATGCTGTATGGGTATGTGCGGCAAAATCTTGTGTTTCCCG 312

RESULT 31
AAC97579
ID AAC97579 standard; DNA; 324 BP.

XX
AC AAC97579;
XX
DT 27-FEB-2001 (first entry)
XX
DE DNA encoding a serine protease inhibitory peptide.
XX
KW Serine protease inhibitor; cytostatic; anti-inflammatory; arthritis;
KW protease mediated tissue destruction; emphysema; glomerulonephritis;
KW periodontitis; muscular dystrophy; tumour invasion; chymotrypsin;
KW elastase; ds.
XX
OS Synthetic.
XX
PN US6132990-A.
XX
PD 17-OCT-2000.
XX
PF 07-JUN-1991; 91US-0712354.
XX
PR 03-JAN-1989; 89US-0293042.
PR 06-DEC-1984; 84US-0678822.
PR 02-DEC-1985; 85US-0803471.
PR 29-JUL-1986; 86US-0890526.
PR 30-MAR-1987; 87US-0031846.
PR 04-AUG-1987; 87US-0082962.
XX
PA (AMGE-) AMGEN BOULDER INC.
XX
PI Bandyopadhyay PK, Eisenberg SP, Stetler GL, Thompson RC;
XX
DR WPI; 2000-678667/66.
XX
PT New serine protease inhibitors and DNA sequences for treating a
PT protease-mediated condition or tissue destruction e.g. emphysema or
PT tumor invasion and for recombinant production of inhibitors -
XX
PS Example 2; Column 30-32; 47pp; English.
XX
CC This invention relates to new purified and isolated mammalian serine
CC protease inhibitor proteins which comprise at least 8 cysteine residues
CC and no more than 107 amino acids. The protease inhibitors are capable of

CC inhibiting chymotrypsin and elastase. Sequences AAC97526 - AAC97534,
CC AAC97574 - AAC97581 and AAC97614 - AAC97641 represent DNA encoding the
CC protease inhibitors of the invention, and include oligonucleotide
CC sequences used in the isolation and characterisation of the proteins.
CC Primers AAC97535 - AAC97573 and AAC95582 - AAC97613 are used in the
CC construction of DNA encoding the protease inhibitors. Peptide sequences
CC AAB53098 - AAB53122 represent the protease inhibitors of the invention and
CC various other peptides used in the isolation of the protease inhibitors.
CC The protease inhibitors have cytostatic and anti-inflammatory activity.
CC The serine protease inhibitor protein is useful for treating a
CC protease-mediated condition, which includes protease mediated tissue
CC destruction, e.g. emphysema, arthritis, glomerulonephritis,
CC periodontitis, muscular dystrophy or tumour invasion. It is also useful
CC for modulating protease activity. The DNAs are useful for producing the
CC proteins, especially by recombinant methods.

XX
SQ Sequence 324 BP; 86 A; 85 C; 87 G; 66 T; 0 other;

Alignment Scores:
Pred. No.: 7.52e-08 Length: 324
Score: 143.00 Matches: 27
Percent Similarity: 55.36% Conservative: 4
Best Local Similarity: 48.21% Mismatches: 23
Query Match: 43.73% Indels: 2
DB: 21 Gaps: 1

09-833799-13B (1-57) x AAC97579 (1-324)

Qy 3 GluProValLysGlyProValSerThr-----LysProGlySerCysProIleLeu 20
AAI67488
Db 145 GATCCGGTTGATACCCCGAACCCGACTCGTCGAAACCCGGTAAATGCCCGTAACCTAT 204

Qy 21 IleArgCysAlaMetLeuAsnProProAsnArgCysLeuLysAspThrAspCysProGly 40
Db 205 GGCCAGTGTCTGATGCTGAACCCCGCGAACTTCTGCGAAATGGACGCCAGTGTAACCGA 264

Qy 41 IleLysLysCysCysGluGlySerCysGlyMetAlaCysPheValPro 56
Db 265 GATCTGAAATGCTGTATGGGTATGTGCGGCAAAATCTTGTGTTTCCCG 312

RESULT 32
AAI67488
ID AAI67488 standard; DNA; 324 BP.

XX
AC AAI67488;
XX
DT 11-FEB-2002 (first entry)
XX
DE Recombinant serine protease inhibitor DNA.
XX
KW Serine protease inhibitor protein; recombinant; leukocyte elastase;
KW trypsin; ds.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT CDS 1..324
FT /*tag= a
FT /product= "serine protease inhibitor"

XX
PN US6291662-B1.
XX
PD 18-SEP-2001.
XX
PF 22-SEP-1998; 98US-0158085.
XX
PR 30-MAR-1987; 87US-0031846.
PR 06-AUG-1990; 90US-0563832.
PR 22-JUL-1994; 94US-0279056.
PR 05-DEC-1984; 84US-0678222.
PR 02-DEC-1985; 85US-0803471.
PR 29-JUL-1986; 86US-0890526.
XX

PA (AMGE-) AMGEN INC.
XX Bandyopadhyay PK, Eisenberg SP, Stetler GL, Thompson RC;
PI WPI; 2001-637974/73.
XX P-PSDB; AAG65993.
DR
XX
PT New DNA sequences, useful in recombinant DNA techniques for directing
PT the production of a serine protease inhibitor protein, e.g. leukocyte
PT elastase or trypsin
XX
PS Example 1; Column 4; 37pp; English.
XX
CC The invention relates to a DNA sequence encoding an analog of a mammalian
CC serine protease inhibitor protein. The analog comprises at least eight
CC cysteine residues and possesses serine protease inhibitor activity, and
CC at least one operational element consisting of a promoter, an operator,
CC a leader sequence, a Shine-Dalgarno sequence, a ribosome binding site
CC or a terminator codon. The DNA is useful in recombinant DNA techniques
CC for directing the production of a serine protease inhibitor protein,
CC e.g. leukocyte elastase or trypsin. The present sequence represents the
CC coding strand of a DNA that directs the manufacture of a recombinant
CC serine protease inhibitor.
XX
SQ Sequence 324 BP; 87 A; 85 C; 88 G; 64 T; 0 other;
Alignment Scores:
Pred. No.: 7.52e-08 Length: 324
Score: 143.00 Matches: 27
Percent Similarity: 55.36% Conservative: 4
Best Local Similarity: 48.21% Mismatches: 23
Query Match: 43.73% Indels: 2
DB: 22 Gaps: 1
09-833799-13B (1-57) x AAI67488 (1-324)
QY 3 GluProValLysGlyProValSerThr-----LysProGlySerCysProIleLeu 20
Db 145 GATCCGGTTGATACCCCGAACCCGACTCGTGGAAACCGGTAATGCCGGTAACCTAT 204
QY 21 IleArgCysAlaMetLeuAsnProProAsnArgCysLeuLysAspThrAspCysProGly 40
Db 205 GGCCAGTGTCTGATGCTGAACCCGCGAACTTCTGCGAAATGGACGGCCAGTGTAAACGA 264
QY 41 IleLysLysCysCysGluGlySerCysGlyMetAlaCysPheValPro 56
Db 265 GATCTGAAATGCTGTATGGTATGTGCGGCAAAATCTTGTGTTTCCCG 312
RESULT 33
AAI67489
ID AAI67489 standard; DNA; 324 BP.
XX
AC AAI67489;
XX
DT 11-FEB-2002 (first entry)
XX
DE Recombinant secretory leukocyte protease inhibitor (SLPI) DNA.
XX
KW Serine protease inhibitor protein; recombinant; leukocyte elastase;
KW trypsin; secretory leukocyte protease inhibitor; SLPI; ds.
XX
OS Homo sapiens.
XX
PN US6291662-B1.
XX
XX 18-SEP-2001.
XX
PF 22-SEP-1998; 98US-0158085.
XX
PR 30-MAR-1987; 87US-0031846.
PR 06-AUG-1990; 90US-0563832.
PR 22-JUL-1994; 94US-0279056.
PR -05-DEC-1984; 84US-0678222.

PR 02-DEC-1985; 85US-0803471.
PR 29-JUL-1986; 86US-0890526.
XX
PA (AMGE-) AMGEN INC.
XX
PI Bandyopadhyay PK, Eisenberg SP, Stetler GL, Thompson RC;
XX
DR WPI; 2001-637974/73.
DR P-PSDB; AAG65996.
XX
PT New DNA sequences, useful in recombinant DNA techniques for directing
PT the production of a serine protease inhibitor protein, e.g. leukocyte
PT elastase or trypsin
XX
XX Disclosure; Column 5; 37pp; English.
PS
XX The invention relates to a DNA sequence encoding an analog of a mammalian
CC serine protease inhibitor protein. The analog comprises at least eight
CC cysteine residues and possesses serine protease inhibitor activity, and
CC at least one operational element consisting of a promoter, an operator,
CC a leader sequence, a Shine-Dalgarno sequence, a ribosome binding site
CC or a terminator codon. The DNA is useful in recombinant DNA techniques
CC for directing the production of a serine protease inhibitor protein,
CC e.g. leukocyte elastase or trypsin. The present sequence represents the
CC coding strand of a DNA that directs the manufacture of a recombinant
CC secretory leukocyte protease inhibitor (SLPI).
XX
SQ Sequence 324 BP; 86 A; 85 C; 87 G; 66 T; 0 other;
Alignment Scores:
Pred. No.: 7.52e-08 Length: 324
Score: 143.00 Matches: 27
Percent Similarity: 55.36% Conservative: 4
Best Local Similarity: 48.21% Mismatches: 23
Query Match: 43.73% Indels: 2
DB: 22 Gaps: 1
09-833799-13B (1-57) x AAI67489 (1-324)
QY 3 GluProValLysGlyProValSerThr-----LysProGlySerCysProIleLeu 20
Db 145 GATCCGGTTGATACCCCGAACCCGACTCGTGGAAACCGGTAATGCCGGTAACCTAT 204
QY 21 IleArgCysAlaMetLeuAsnProProAsnArgCysLeuLysAspThrAspCysProGly 40
Db 205 GGCCAGTGTCTGATGCTGAACCCGCGAACTTCTGCGAAATGGACGGCCAGTGTAAACGA 264
QY 41 IleLysLysCysCysGluGlySerCysGlyMetAlaCysPheValPro 56
Db 265 GATCTGAAATGCTGTATGGTATGTGCGGCAAAATCTTGTGTTTCCCG 312
RESULT 34
ABA83108
ID ABA83108 standard; DNA; 399 BP.
XX
AC ABA83108;
XX
DT 08-FEB-2002 (first entry)
XX
DE Secretory leukocyte protease inhibitor ovarian tumour marker gene, #53.
XX
KW Ovarian tumour marker gene; human; overexpression; upregulation;
KW epithelial tumour; cancer; diagnosis; prognosis; disease monitoring;
KW identification; serous cystadenoma; borderline serous tumour;
KW serous cystadenocarcinoma; mucinous cystadenocarcinoma;
KW mucinous cystadenoma; borderline mucinous tumour; endometrioid carcinoma;
KW undifferentiated carcinoma; clear cell adenocarcinoma; cystadenofibroma;
KW adenofibroma; Brenner tumour; serial analysis of gene expression; SAGE;
KW immune response pathway; cell proliferation regulation; protein folding;
KW membrane localised; secreted; therapeutic target; cytostatic;
KW gene therapy; vaccine; ds.
XX
OS Homo sapiens.

XX WO200175177-A2.
PN 11-OCT-2001.
XX 03-APR-2001; 2001WO-US10947.
XX 03-APR-2000; 2000US-194336P.
XX (USSH) US DEPT HEALTH & HUMAN SERVICES.
PA Morin PJ, Sherman-Baust CA, Pizer ES, Hough CD;
XX WPI; 2001-626450/72.
DR P-PSDB; ABB50282.
XX
PT Detecting and identifying ovarian tumor, identifying increased risk for
PT developing ovarian cancer, and determining effectiveness of ovarian
PT cancer treatment, by measuring expression level of ovarian tumor marker
PT gene -
XX
PS Claim 23; Page 98; 140pp; English.
XX
CC The invention relates to methods for diagnosing and prognosing ovarian
CC tumours in an individual via the detection and measurement of the
CC expression of ovarian tumour marker genes (ABA83081-ABA83122, ABA83180,
CC ABA83182 and ABA83184) or segments thereof (ABA83123-ABA83169, ABA83179,
CC ABA83181 and ABA83183). The methods of the invention are useful for
CC detecting an ovarian tumour in a patient, for identifying an individual
CC at increased risk for developing ovarian cancer, in prognostic tests for
CC assessing the relative severity of ovarian cancer, in tests for
CC monitoring a patient in remission from ovarian cancer and in tests for
CC monitoring disease status in a patient being treated for ovarian cancer.
CC The methods can additionally be used to identify a particular tumour as
CC being an ovarian tumour (i.e., an epithelial ovarian tumour selected from
CC serous cystadenoma, borderline serous tumour, serous cystadenocarcinoma,
CC mucinous cystadenoma, borderline mucinous tumour, mucinous
CC cystadenocarcinoma, endometrioid carcinoma, undifferentiated carcinoma,
CC clear cell adenocarcinoma, cystadenofibroma, adenofibroma and Brenner
CC tumour. The ovarian tumour marker genes of the invention were identified
CC using SAGE (serial analysis of gene expression) and were found to be
CC overexpressed in a broad variety of ovarian epithelial tumour cells
CC relative to normal ovarian epithelial cells. The marker genes are
CC implicated in immune response pathways, in the regulation of cell
CC proliferation and in protein folding, and many of these are membrane-
CC localised or secreted. In addition to their use as diagnostic and
CC prognostic markers, the ovarian tumour marker genes or their encoded
CC proteins may be used as therapeutic targets for the treatment and
CC prevention of ovarian cancer. Sequences ABA83081-ABA83122, ABA83180,
CC ABA83182 and ABA83184 represent the ovarian tumour marker genes of
CC the invention.
XX
SQ Sequence 399 BP; 91 A; 100 C; 111 G; 97 T; 0 other;

Alignment Scores:
Pred. No.: 9.7e-08 Length: 399
Score: 143.00 Matches: 27
Percent Similarity: 55.36% Conservative: 4
Best Local Similarity: 48.21% Mismatches: 23
Query Match: 43.73% Indels: 2
DB: 22 Gaps: 1

09-833799-13B (1-57) x ABA83108 (1-399)

QY 3 GluProValLysGlyProValSerThr-----LysProGlySerCysProIleIleLeu 20
:::|||||
Db 220 GATCCTGTTGACACCCCAACCAACAGGAGGAGCCTGGGAAGTCCCCAGTGACTTAT 279
:::|||||
QY 21 IleArgCysAlaMetLeuAsnProProAsnArgCysLeuLysAspThrAspCysProGly 40
:::|||||
Db 280 GGCCAATGTTTGATGCTTAACCCCCCAATTTCTGTGAGATGGATGGCCAGTCAAGCGT 339
:::|||||
QY 41 IleLysLysCysCysGluGlySerCysGlyMetAlaCysPheValPro 56
:::|||||

Db 340 GACTTGAAGTGTTCATGGGCATGTGTGGAAATCTCGCTTCCCT 387
|||||
RESULT 35
AAX16272
ID AAX16272 standard; DNA; 420 BP.
XX
AC AAX16272;
XX
DT 19-APR-1999 (first entry)
XX
DE DNA sequence of ompA-tc-met-SLPI.
XX
KW Serine protease inhibitor; human; emphysema; arthritis; peridontitis;
KW muscular dystrophy; tumour invasion; glomerulonephritis; sepsis;
KW acute leukemia; ds.
XX
OS Homo sapiens.
OS Synthetic.
XX
PN US5871956-A.
XX
PD 16-FEB-1999.
XX
XX 22-JUL-1994; 94US-0279056.
PF
XX 30-MAR-1987; 87US-0031846.
PR 05-DEC-1984; 84US-0678822.
PR 29-JUL-1986; 86US-0890526.
PR 03-SEP-1986; 86US-0903471.
PR 06-AUG-1990; 90US-0563832.
PR 22-JUL-1994; 94US-0279056.
XX
PA (AMGE-) AMGEN INC.
XX
PI Bandyopadhyay PK, Eisenberg SP, Stetler GL, Thompson RC;
XX WPI; 1999-166640/14.
DR
XX New DNA sequence encoding mammalian serine protease inhibitor -
XX useful for recombinantly producing inhibitors with different
XX specificities and treating diseases such as emphysema, arthritis,
XX muscular dystrophy, and tumour invasion
XX
PS Example 3; Column 37-40; 37pp; English.
XX
CC The present invention describes a DNA sequence (A) encoding an analogue
CC of a mammalian serine protease inhibitor (B). The DNA sequences and
CC recombinant methods allow manufacture of a class of inhibitors of e.g.
CC cathepsin G, elastase, and trypsin, with different specificities. The
CC recombinant serine protease product can be directed to act
CC intracellularly or extracellularly and is useful in treating conditions
CC caused by a disturbance in the native protease/protease inhibitor
CC balance such as emphysema, arthritis, glomerulonephritis, peridontitis,
CC muscular dystrophy, tumour invasion, sepsis, and acute leukemia. (A)
CC allows (B) to be recombinantly produced in sufficient quantities and
CC purities so as to provide economical pharmaceutical compositions. (B) is
CC resistant to heat, acid, and a variety of proteolytic enzymes, is
CC thermodynamically stable in extracellular conditions, and exhibits a
CC high degree of self assembly forming an active tertiary structure in the
CC absence of biochemical stimuli. The present sequence represents the DNA
CC sequence of ompA-tc-met-SLPI.
XX
SQ Sequence 420 BP; 123 A; 98 C; 112 G; 87 T; 0 other;

Alignment Scores:
Pred. No.: 1.03e-07 Length: 420
Score: 143.00 Matches: 27
Percent Similarity: 55.36% Conservative: 4
Best Local Similarity: 48.21% Mismatches: 23
Query Match: 43.73% Indels: 2
DB: 20 Gaps: 1

CC e.g. leukocyte elastase or trypsin. The present sequence represents the
CC DNA sequence coding for ompA-tc-met-secretory leukocyte protease
CC inhibitor (SLPI).

XX Sequence 436 BP; 128 A; 102 C; 116 G; 90 T; 0 other;

Alignment Scores:

Pred. No.: 1.08e-07 Length: 436
Score: 143.00 Matches: 27
Percent Similarity: 55.36% Conservative: 4
Best Local Similarity: 48.21% Mismatches: 23
Query Match: 43.73% Indels: 2
DB: 22 Gaps: 1

09-833799-13B (1-57) x AAI67583 (1-436)

QY 3 GluProValLysGlyProValSerThr-----LysProGlySerCysProIleLeu 20
Db 245 GATCCGGTTGATACCCGAAACCCGACTCGTCGAAACCGGGTAAATGCCCGTAACCTAT 304
QY 21 IleArgCysAlaMetLeuAsnProProAsnArgCysLeuLysAspThrAspCysProGly 40
Db 305 GGCCAGTGTCTGATGCTGAACCCCGCGAACTTCTCGAAATGGACGCCAGTGTAACCGA 364
QY 41 IleLysLysCysCysGluGlySerCysGlyMetAlaCysPheValPro 56
Db 365 GATCTGAAATGCTGTATGGGTATGTGCGCAAAATCTTGTGTTTCCCCG 412

RESULT 38

AAX16271

ID AAX16271 standard; DNA; 460 BP.

XX AAX16271;

DT 19-APR-1999 (first entry)

XX DNA sequence of ompA-SLPI.

KW Serine protease inhibitor; human; emphysema; arthritis; peridontitis;
KW muscular dystrophy; tumour invasion; glomerulonephritis; sepsis;
KW acute leukemia; ds.

OS Homo sapiens.

XX Synthetic.

XX US5871956-A.

PD 16-FEB-1999.

XX 22-JUL-1994; 94US-0279056.

PR 30-MAR-1987; 87US-0031846.

PR 05-DEC-1984; 84US-0678822.

PR 29-JUL-1986; 86US-0890526.

PR 03-SEP-1986; 86US-0903471.

PR 06-AUG-1990; 90US-0563832.

XX 22-JUL-1994; 94US-0279056.

PA (AMGE-) AMGEN INC.

PI Bandyopadhyay PK, Eisenberg SP, Stetler GL, Thompson RC;

XX WPI; 1999-166640/14.

XX New DNA sequence encoding mammalian serine protease inhibitor -
PT useful for recombinantly producing inhibitors with different
PT specificities and treating diseases such as emphysema, arthritis,
PT muscular dystrophy, and tumour invasion

PS Example 3; Column 37-38; 37pp; English.

XX The present invention describes a DNA sequence (A) encoding an analogue
CC of a mammalian serine protease inhibitor (B). The DNA sequences and

CC recombinant methods allow manufacture of a class of inhibitors of e.g.
CC cathepsin G, elastase, and trypsin, with different specificities. The
CC recombinant serine protease product can be directed to act
CC intracellularly or extracellularly and is useful in treating conditions
CC caused by a disturbance in the native protease/protease inhibitor
CC balance such as emphysema, arthritis, glomerulonephritis, peridontitis,
CC muscular dystrophy, tumour invasion, sepsis, and acute leukemia. (A)
CC allows (B) to be recombinantly produced in sufficient quantities and
CC purities so as to provide economical pharmaceutical compositions. (B) is
CC resistant to heat, acid, and a variety of proteolytic enzymes, is
CC thermodynamically stable in extracellular conditions, and exhibits a
CC high degree of self assembly forming an active tertiary structure in the
CC absence of biochemical stimuli. The present sequence represents the DNA
CC sequence of ompA-SLPI.

XX Sequence 460 BP; 123 A; 114 C; 125 G; 98 T; 0 other;

Alignment Scores:

Pred. No.: 1.16e-07 Length: 460
Score: 143.00 Matches: 27
Percent Similarity: 55.36% Conservative: 4
Best Local Similarity: 48.21% Mismatches: 23
Query Match: 43.73% Indels: 2
DB: 20 Gaps: 1

09-833799-13B (1-57) x AAX16271 (1-460)

QY 3 GluProValLysGlyProValSerThr-----LysProGlySerCysProIleLeu 20
Db 269 GATCCGGTTGATACCCGAAACCCGACTCGTCGAAACCGGGTAAATGCCCGTAACCTAT 328

QY 21 IleArgCysAlaMetLeuAsnProProAsnArgCysLeuLysAspThrAspCysProGly 40

Db 329 GGCCAGTGTCTGATGCTGAACCCCGCGAACTTCTCGAAATGGACGCCAGTGTAACCGA 388

QY 41 IleLysLysCysCysGluGlySerCysGlyMetAlaCysPheValPro 56

Db 389 GATCTGAAATGCTGTATGGGTATGTGCGCAAAATCTTGTGTTTCCCCG 436

RESULT 39

AAC97622

ID AAC97622 standard; DNA; 460 BP.

XX AAC97622;

DT 27-FEB-2001 (first entry)

XX DNA encoding OmpA secretory leukocyte protease inhibitor.

KW Serine protease inhibitor; cytostatic; anti-inflammatory; arthritis;
KW protease mediated tissue destruction; emphysema; glomerulonephritis;
KW peridontitis; muscular dystrophy; tumour invasion; chymotrypsin;
KW elastase; ds.

OS Synthetic.

XX US6132990-A.

PD 17-OCT-2000.

XX 07-JUN-1991; 91US-0712354.

PR 03-JAN-1989; 89US-0293042.

PR 06-DEC-1984; 84US-0678822.

PR 02-DEC-1985; 85US-0803471.

PR 29-JUL-1986; 86US-0890526.

PR 30-MAR-1987; 87US-0031846.

PR 04-AUG-1987; 87US-0082962.

XX (AMGE-) AMGEN BOULDER INC.

XX Bandyopadhyay PK, Eisenberg SP, Stetler GL, Thompson RC;

DR WPI; 2000-678667/66.
XX New serine protease inhibitors and DNA sequences for treating a
PT protease-mediated condition or tissue destruction e.g. emphysema or
PT tumor invasion and for recombinant production of inhibitors
XX
XX
PS Example 3; Column 37-38; 47pp; English.
XX
CC This invention relates to new purified and isolated mammalian serine
CC protease inhibitor proteins which comprise at least 8 cysteine residues
CC and no more than 107 amino acids. The protease inhibitors are capable of
CC inhibiting chymotrypsin and elastase. Sequences AAC97526 - AAC97534,
CC AAC97574 - AAC97581 and AAC97614 - AAC97641 represent DNA encoding the
CC protease inhibitors of the invention, and include oligonucleotide
CC sequences used in the isolation and characterisation of the proteins.
CC Primers AAC97535 - AAC97573 and AAC95582 - AAC97613 are used in the
CC construction of DNA encoding the protease inhibitors. Peptide sequences
CC AAB53098 - AAB53122 represent the protease inhibitor of the invention and
CC various other peptides used in the isolation of the protease inhibitors.
CC The protease inhibitors have cytostatic and anti-inflammatory activity.
CC The serine protease inhibitor protein is useful for treating a
CC protease-mediated condition, which includes protease mediated tissue
CC destruction, e.g. emphysema, arthritis, glomerulonephritis,
CC periodontitis, muscular dystrophy or tumour invasion. It is also useful
CC for modulating protease activity. The DNAs are useful for producing the
CC proteins, especially by recombinant methods.
XX
SQ Sequence 460 BP; 123 A; 114 C; 125 G; 98 T; 0 other;

Alignment Scores:
Pred. No.: 1.16e-07 Length: 460
Score: 143.00 Matches: 27
Percent Similarity: 55.36% Conservative: 4
Best Local Similarity: 48.21% Mismatches: 23
Query Match: 43.73% Indels: 2
DB: 21 Gaps: 1

09-833799-13B (1-57) x AAC97622 (1-460)

QY 3 GluProValLysGlyProValSerThr-----LysProGlySerCysProIleLeu 20
Db 269 GATCCGGTTGATACCCGAAACCGGTAATGCGGTAACCTAT 328
QY 21 IleArgCysAlaMetLeuAsnProProAsnArgCysLeuLysAspThrAspCysProGly 40
Db 329 GGCCAGTGTCTGATGCTGAACCCGCGAACTTCTGCGAAATGGACGCCAGTGTAACCGA 388
QY 41 IleLysLysCysCysGluGlySerCysGlyMetAlaCysPheValPro 56
Db 389 GATCTGAAATGCTGATGGGTATGCGGCAAAATCTTGTGTTTCCCG 436

RESULT 40

AAI67582
ID AAI67582 standard; DNA; 460 BP.
XX
AC AAI67582;
XX
DT 11-FEB-2002 (first entry)
XX
DE DNA sequence coding for ompA SLPI.

XX Serine protease inhibitor protein; recombinant; leukocyte elastase;
KW trypsin; secretory leukocyte protease inhibitor; SLPI; ds.
XX

OS Homo sapiens.
XX
PN US6291662-B1.
XX
PD 18-SEP-2001.
XX
PF 22-SEP-1998; 98US-0158085.
XX
PR 30-MAR-1987; 87US-0031846.

PR 06-AUG-1990; 90US-0563832.
PR 22-JUL-1994; 94US-0279056.
PR 05-DEC-1984; 84US-0678222.
PR 02-DEC-1985; 85US-0803471.
PR 29-JUL-1986; 86US-0890526.
XX
XX (AMGE-) AMGEN INC.
XX
PI Bandyopadhyay PK, Eisenberg SP, Stetler GL, Thompson RC;
XX
DR WPI; 2001-637974/73.

XX New DNA sequences, useful in recombinant DNA techniques for directing
PT the production of a serine protease inhibitor protein, e.g. leukocyte
PT elastase or trypsin
XX

PS Example 3; Column 36; 37pp; English.

XX The invention relates to a DNA sequence encoding an analog of a mammalian
CC serine protease inhibitor protein. The analog comprises at least eight
CC cysteine residues and possesses serine protease inhibitor activity, and
CC at least one operational element consisting of a promoter, an operator,
CC a leader sequence, a Shine-Dalgarno sequence, a ribosome binding site
CC or a terminator codon. The DNA is useful in recombinant DNA techniques
CC for directing the production of a serine protease inhibitor protein,
CC e.g. leukocyte elastase or trypsin. The present sequence represents the
CC DNA sequence coding for ompA-secretory leukocyte protease inhibitor
CC (SLPI).

SQ Sequence 460 BP; 123 A; 114 C; 125 G; 98 T; 0 other;

Alignment Scores:
Pred. No.: 1.16e-07 Length: 460
Score: 143.00 Matches: 27
Percent Similarity: 55.36% Conservative: 4
Best Local Similarity: 48.21% Mismatches: 23
Query Match: 43.73% Indels: 2
DB: 22 Gaps: 1

09-833799-13B (1-57) x AAI67582 (1-460)

QY 3 GluProValLysGlyProValSerThr-----LysProGlySerCysProIleLeu 20
Db 269 GATCCGGTTGATACCCGAAACCGGTAATGCGGTAACCTAT 328
QY 21 IleArgCysAlaMetLeuAsnProProAsnArgCysLeuLysAspThrAspCysProGly 40
Db 329 GGCCAGTGTCTGATGCTGAACCCGCGAACTTCTGCGAAATGGACGCCAGTGTAACCGA 388
QY 41 IleLysLysCysCysGluGlySerCysGlyMetAlaCysPheValPro 56
Db 389 GATCTGAAATGCTGATGGGTATGCGGCAAAATCTTGTGTTTCCCG 436

RESULT 41

AAAN70929
ID AAAN70929 standard; cDNA; 498 BP.
XX
AC AAAN70929;
XX
DT 21-FEB-1991 (first entry)
XX

XX Sequence encoding protein with the biological activity of HUSI
DE (human seminal plasma inhibitor) type I inhibitors encoded on pRH31.
XX

KW Bronchitis therapy; cervix inflammation; hyperfibrinolysis; ds.

XX Homo sapiens.

XX
FH Key Location/Qualifiers
FT CDS 25..297
FT /*tag= a

PN DE3600571-A.

XX 06-AUG-1987.
PD 10-JAN-1986; 86DE-3600571.
XX 10-JAN-1986; 86DE-3600571.
PF 10-JAN-1986; 86DE-3600571.
XX 10-JAN-1986; 86DE-3600571.
PR (CHEF) GRUNENTHAL GMBH.
XX
XX Heinzl R, Appelhans H, Gassen HG, Machleidt W;
PI WPI; 1987-222038/32.
XX P-PSDB; AAP70583.
DR
XX New protein with human seminal plasma inhibitor activity - and
PT new DNA coding sequences, vectors and transformed organisms,
PT useful e.g. for treating bronchitis or inflammation
XX
XX Claim 4; Fig 4; 28pp; German.
PS
XX HUSI-type inhibitors are useful for treating chronic bronchitis,
CC chronic cervical inflammation; inflammation associated with
CC excessive mucus prodn.; post-operative bleeding caused by
CC hyperfibrinolysis, or shock. They are esp. suitable for
CC administration as spray or inhalation. The expression control
CC sequence on the DNA fragment in AAN70928 is bound to the 5'-end of
CC HUSI cDNA in an expression vector.
XX
SQ Sequence 498 BP; 110 A; 131 C; 141 G; 116 T; 0 other;

Alignment Scores:
Pred. No.: 1.27e-07 Length: 498
Score: 143.00 Matches: 27
Percent Similarity: 55.36% Conservative: 4
Best Local Similarity: 48.21% Mismatches: 23
Query Match: 43.73% Indels: 2
DB: 8 Gaps: 1

09-833799-13B (1-57) x AAN70929 (1-498)
QY 3 GluProValLysGlyProValSerThr-----LysProGlySerCysProIleIleLeu 20
Db 118 GATCCTGTTGACACCCCAACCAAGAGGAGGAGCCTGGGAAGTGGCCAGTGACTTAT 177
QY 21 IleArgCysAlaMetLeuAsnProAsnArgCysLeuLysAspThrAspCysProGly 40
Db 178 GGCCAATGTTGATGCTTAACCCCAATTTCTGTGAGATGGATGGCCAGTGCAAGCGT 237
QY 41 IleLysLysCysCysGluGlySerCysGlyMetAlaCysPheValPro 56
Db 238 GACTTGAAGTGTTCATGGGCATGTGTGGGAAATCCTGCGTTTCCCT 285

RESULT 42
AAN70930
ID AAN70930 standard; cDNA; 565 BP.
XX
AC AAN70930;
XX
DT 21-MAY-1991 (first entry)
XX
DE Sequence encoding protein with the biological activity of HUSI
DE (human seminal plasma inhibitor) type I inhibitors encoded on pRH 34.
XX
KW Bronchitis therapy; cervix inflammation; hyperfibrinolysis; ds.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT CDS 59..457
FT /*tag= a
XX
PN DE3600571-A.
XX

PD 06-AUG-1987.
XX 10-JAN-1986; 86DE-3600571.
PF 10-JAN-1986; 86DE-3600571.
PR 10-JAN-1986; 86DE-3600571.
XX (CHEF) GRUNENTHAL GMBH.
XX
XX Heinzl R, Appelhans H, Gassen HG, Machleidt W, Seemuller U;
PI WPI; 1987-222038/32.
XX P-PSDB; AAP70584.
DR
XX New protein with human seminal plasma inhibitor activity - and
PT new DNA coding sequences, vectors and transformed organisms,
PT useful e.g. for treating bronchitis or inflammation
XX
XX Claim 5; Fig 5; 28pp; German.
PS
XX HUSI-type I inhibitors are useful for treating chronic bronchitis,
CC chronic cervical inflammation; inflammation associated with
CC excessive mucus prodn.; post-operative bleeding caused by
CC hyperfibrinolysis, and shock. They are esp. suitable for
CC administration as spray or inhalation. The expression control
CC sequence on the DNA fragment in AAN70928 is bound to the 5' end of
CC HUSI cDNA in an expression vector.
XX
SQ Sequence 565 BP; 106 A; 162 C; 171 G; 126 T; 0 other;

Alignment Scores:
Pred. No.: 1.49e-07 Length: 565
Score: 143.00 Matches: 27
Percent Similarity: 55.36% Conservative: 4
Best Local Similarity: 48.21% Mismatches: 23
Query Match: 43.73% Indels: 2
DB: 8 Gaps: 1

09-833799-13B (1-57) x AAN70930 (1-565)
QY 3 GluProValLysGlyProValSerThr-----LysProGlySerCysProIleIleLeu 20
Db 278 GATCCTGTTGACACCCCAACCAAGAGGAGGAGCCTGGGAAGTGGCCAGTGACTTAT 337
QY 21 IleArgCysAlaMetLeuAsnProAsnArgCysLeuLysAspThrAspCysProGly 40
Db 338 GGCCAATGTTGATGCTTAACCCCAATTTCTGTGAGATGGATGGCCAGTGCAAGCGT 397
QY 41 IleLysLysCysCysGluGlySerCysGlyMetAlaCysPheValPro 56
Db 398 GACTTGAAGTGTTCATGGGCATGTGTGGGAAATCCTGCGTTTCCCT 445

RESULT 43
AAF32043
ID AAF32043 standard; cDNA; 565 BP.
XX
AC AAF32043;
XX
DT 10-APR-2001 (first entry)
XX
DE Human antileukoprotease coding sequence.
XX
KW Human; cytostatic; gynecological; elastase; cathepsin G; gene therapy;
KW endometriosis; antileukoprotease; ALP; secretory leukoprotease inhibitor;
KW SLPI; alphas-1-proteinase inhibitor; alphas-1-PI; ss.
XX
OS Homo sapiens.
XX
PN WO200101998-A2.
XX
PD 11-JAN-2001.
XX
PF 29-JUN-2000; 2000WO-IB000990.
XX

PR	05-JUN-2000;	2000US-209531P.
PR	18-SEP-2000;	2000US-233133P.
PR	18-SEP-2000;	2000US-233617P.
PR	20-SEP-2000;	2000US-234009P.
PR	20-SEP-2000;	2000US-234034P.
PR	20-SEP-2000;	2000US-234052P.
PR	22-SEP-2000;	2000US-234509P.
PR	22-SEP-2000;	2000US-234567P.
PR	25-SEP-2000;	2000US-234923P.
PR	25-SEP-2000;	2000US-234924P.
PR	25-SEP-2000;	2000US-235077P.
PR	25-SEP-2000;	2000US-235082P.
PR	25-SEP-2000;	2000US-235134P.
PR	25-SEP-2000;	2000US-235280P.
PR	26-SEP-2000;	2000US-235637P.
PR	26-SEP-2000;	2000US-235638P.
PR	27-SEP-2000;	2000US-235711P.
PR	27-SEP-2000;	2000US-235720P.
PR	27-SEP-2000;	2000US-235840P.
PR	27-SEP-2000;	2000US-235863P.
PR	28-SEP-2000;	2000US-236028P.
PR	28-SEP-2000;	2000US-236032P.
PR	28-SEP-2000;	2000US-236033P.
PR	28-SEP-2000;	2000US-236034P.
PR	28-SEP-2000;	2000US-236109P.
PR	28-SEP-2000;	2000US-236111P.
PR	29-SEP-2000;	2000US-236842P.
PR	29-SEP-2000;	2000US-236891P.
PR	02-OCT-2000;	2000US-237172P.
PR	02-OCT-2000;	2000US-237173P.
PR	02-OCT-2000;	2000US-237278P.
PR	02-OCT-2000;	2000US-237294P.
PR	02-OCT-2000;	2000US-237295P.
PR	02-OCT-2000;	2000US-237316P.
PR	03-OCT-2000;	2000US-237425P.
PR	03-OCT-2000;	2000US-237598P.
PR	03-OCT-2000;	2000US-237604P.
PR	03-OCT-2000;	2000US-237606P.
PR	03-OCT-2000;	2000US-237608P.
PR	01-NOV-2000;	2000US-244867P.
PR	01-NOV-2000;	2000US-245084P.
XX		
PA	(AVAL-) AVALON PHARM.	
XX		
PI	Young PE, Augustus M, Cart	
PI	Soppet DR, Weaver Z;	
XX		
DR	WPI; 2002-188264/24.	
XX		
PT	Screening for anti-neoplastic	
PT	chemical agent to be tested i	
PT	determining a change in expre	
XX		
PS	Claim 1; SEQ ID 5016; 44pp; E	
XX		
CC	The present invention describ	
CC	anti-neoplastic agent. The me	
CC	agent to be tested for anti-r	
CC	expression of at least one ge	
CC	comprises a sequence (S) sele	
CC	to ABL70110), or is at least	
CC	expression is indicative of a	
CC	activity and can be used in g	
CC	an anti-neoplastic agent, and	
CC	is the data collected with re	
CC	result of M1, and the data is	
CC	structure and/or properties c	
CC	treatment of cancer such as c	
CC	oesophageal, ovarian, kidney,	
CC	adenocarcinoma, carcinoma, cl	
CC	infiltrating lobular cancer,	
CC	carcinoma, papillary carcinom	
XX		

SQ Sequence 594 BP; 132 A; 156 C; 155 G; 151 T; 0 other;

Alignment Scores:
Pred. No.: 1.58e-07 Length: 594
Score: 143.00 Matches: 27
Percent Similarity: 55.36% Conservative: 4
Best Local Similarity: 48.21% Mismatches: 23
Query Match: 43.73% Indels: 2
DB: 24 Gaps: 1

09-833799-13B (1-57) x ABL66679 (1-594)

QY 3 GluProValLysGlyProValSerThr-----LysProGlySerCysProIleLeu 20
Db 238 GATCCTGTTGACACCCCAACCAAGGAGGAGCTGGGAAGTGGCCAGTGAATTAT 297
QY 21 IleArgCysAlaMetLeuAsnProProAsnArgCysLeuLysAspThrAspCysProGly 40
Db 298 GGCAATGTTGATGCTTAACCCCAATTTCTGTGAGATGGATGGCCAGTGAAGCGT 357
QY 41 IleLysLysCysCysGluGlySerCysGlyMetAlaCysPheValPro 56
Db 358 GACTTGAAGTGTTCATGGGCATGTGTGGAAATCTCGGTTTCCCT 405

RESULT 45
ABL67548
ID ABL67548 standard; DNA; 594 BP.
XX AC ABL67548;
XX 15-MAY-2002 (first entry)
DT Thyroid cancer related gene sequence SEQ ID NO:5885.

XX Human; cancer; colon; breast; ovary; oesophagus; kidney; thyroid;
KW stomach; lung; prostate; pancreas; carcinoma; antitumour; cancerous;
KW cytostatic; gene therapy; antineoplastic; Wilm's tumour; adenocarcinoma;
KW gene; ds.

XX Homo sapiens.
XX WO200194629-A2.
XX 13-DEC-2001.
XX 30-MAY-2001; 2001WO-US10838.

PR 05-JUN-2000; 2000US-209473P.
PR 05-JUN-2000; 2000US-209531P.
PR 18-SEP-2000; 2000US-233133P.
PR 18-SEP-2000; 2000US-233617P.
PR 20-SEP-2000; 2000US-234009P.
PR 20-SEP-2000; 2000US-234034P.
PR 20-SEP-2000; 2000US-234052P.
PR 22-SEP-2000; 2000US-234509P.
PR 22-SEP-2000; 2000US-234567P.
PR 25-SEP-2000; 2000US-234923P.
PR 25-SEP-2000; 2000US-234924P.
PR 25-SEP-2000; 2000US-235077P.
PR 25-SEP-2000; 2000US-235082P.
PR 25-SEP-2000; 2000US-235134P.
PR 25-SEP-2000; 2000US-235280P.
PR 26-SEP-2000; 2000US-235637P.
PR 26-SEP-2000; 2000US-235638P.
PR 27-SEP-2000; 2000US-235711P.
PR 27-SEP-2000; 2000US-235720P.
PR 27-SEP-2000; 2000US-235840P.
PR 27-SEP-2000; 2000US-235863P.
PR 28-SEP-2000; 2000US-236028P.
PR 28-SEP-2000; 2000US-236032P.
PR 28-SEP-2000; 2000US-236033P.
PR 28-SEP-2000; 2000US-236034P.
PR 28-SEP-2000; 2000US-236109P.

PR 28-SEP-2000; 2000US-236111P.
PR 29-SEP-2000; 2000US-236842P.
PR 29-SEP-2000; 2000US-236891P.
PR 02-OCT-2000; 2000US-237172P.
PR 02-OCT-2000; 2000US-237173P.
PR 02-OCT-2000; 2000US-237278P.
PR 02-OCT-2000; 2000US-237294P.
PR 02-OCT-2000; 2000US-237295P.
PR 02-OCT-2000; 2000US-237316P.
PR 03-OCT-2000; 2000US-237425P.
PR 03-OCT-2000; 2000US-237598P.
PR 03-OCT-2000; 2000US-237604P.
PR 03-OCT-2000; 2000US-237606P.
PR 03-OCT-2000; 2000US-237608P.
PR 01-NOV-2000; 2000US-244867P.
PR 01-NOV-2000; 2000US-245084P.
XX (AVAL-) AVALON PHARM.
PA Young PE, Augustus M, Carter KC, Ebner R, Endress G, Horrigan S;
XX Soppet DR, Weaver Z;
PI WPI; 2002-188264/24.
PT Screening for anti-neoplastic agent involves exposing cells to a
PT chemical agent to be tested for anti-neoplastic activity, and
PT determining a change in expression of a gene of a signature gene set -
XX Claim 1; SEQ ID 5885; 44pp; English.

XX The present invention describes a method (M1) for screening for an
CC anti-neoplastic agent. The method involves exposing cells to a chemical
CC agent to be tested for anti-neoplastic activity, determining a change in
CC expression of at least one gene (I) of a signature gene set, where (I)
CC comprises a sequence (S) selected from 8447 sequences (given in ABL61664
CC to ABL70110), or is at least 95% identical to (S), where a change in
CC expression is indicative of anti-neoplastic activity. (I) has cytostatic
CC activity and can be used in gene therapy. M1 can be used for screening
CC an anti-neoplastic agent, and can be used for producing a product which
CC is the data collected with respect to the anti-neoplastic agent as a
CC result of M1, and the data is sufficient to convey the chemical
CC structure and/or properties of the agent. M1 can be used in the
CC treatment of cancer such as colon, breast, stomach, lung, thyroid,
CC oesophageal, ovarian, kidney, prostate or pancreatic cancer.
CC adenocarcinoma, carcinoma, clear cell cancer, infiltrating ductal cancer,
CC infiltrating lobular cancer, squamous cell carcinoma, neuroendocrine
CC carcinoma, papillary carcinoma and Wilm's tumour.

XX Sequence 594 BP; 132 A; 156 C; 155 G; 151 T; 0 other;

Alignment Scores:
Pred. No.: 1.58e-07 Length: 594
Score: 143.00 Matches: 27
Percent Similarity: 55.36% Conservative: 4
Best Local Similarity: 48.21% Mismatches: 23
Query Match: 43.73% Indels: 2
DB: 24 Gaps: 1

09-833799-13B (1-57) x ABL67548 (1-594)

QY 3 GluProValLysGlyProValSerThr-----LysProGlySerCysProIleLeu 20
Db 238 GATCCTGTTGACACCCCAACCAAGGAGGAGCTGGGAAGTGGCCAGTGAATTAT 297
QY 21 IleArgCysAlaMetLeuAsnProProAsnArgCysLeuLysAspThrAspCysProGly 40
Db 298 GGCAATGTTGATGCTTAACCCCAATTTCTGTGAGATGGATGGCCAGTGAAGCGT 357
QY 41 IleLysLysCysCysGluGlySerCysGlyMetAlaCysPheValPro 56
Db 358 GACTTGAAGTGTTCATGGGCATGTGTGGAAATCTCGGTTTCCCT 405

RESULT 46

AAH57161 to AAH57576 represent cell and tissue specific polynucleotide sequences (I). (I) can have cytostatic, immunomodulatory and neuroprotective activities, and can be used in gene therapy. (I) and proteins (II) encoded by then are used in high throughput screening assays to select DNA molecules, RNA molecules, peptide nucleic acids, mimetics, peptides, proteins, agonists, antagonists, antibodies or their fragments, immunoglobulins, inhibitors, drug compounds and pharmaceutical agents. Expression of (I) in a sample indicates the

PD 27-JUN-2002.
XX 18-DEC-2001; 2001WO-US49256.
PF 18-DEC-2000; 2000US-256699P.
XX 20-NOV-2001; 2001US-331966P.
PR (ARRI-) ARRIVA PHARM INC.
XX Barr PJ, Gibson HL, Pemberton P;
PI WPI; 2002-500631/53.
XX P-PSDB; AAU99881.
DR Novel fusion protein useful for inhibiting protease activity associated
XX with a disorder such as emphysema, asthma, comprises a first protease
PT inhibitor comprising alpha 1-antitrypsin and a second protease
PT inhibitor -
XX
PS Example 1; Page 73-73; 134pp; English.
XX This invention relates to a novel fusion protein comprising a first
CC protease inhibitor comprising an alpha1-antitrypsin or its functionally
CC active portion and a second protease inhibitor or its functionally
CC active protein. The fusion proteins of the invention may act as an
CC inhibitor of protease activity. The fusion protein of the invention
CC is useful for inhibiting protease activity associated with a disorder
CC such as emphysema, asthma, chronic obstructive pulmonary disease,
CC cystic fibrosis, otitis media, otitis externa or HIV infection, or
CC for treating an individual suffering from or at risk for a disease or
CC disorder involving unwanted protease activity. The proteins are useful
CC for treating dermatological diseases such as atopic dermatitis, eczema
CC and psoriasis, in inflammatory responses to viral infection, and for
CC treating herpes infection, corneal or epidermal ulceration, chronic
CC non-healing wounds, sepsis, rheumatoid arthritis, periodontal disease,
CC tumour metastasis and tumour angiogenesis, gastric ulceration,
CC osteoporosis, Paget's disease, glomerulonephritis, scleroderma, malaria,
CC bacterial infection, Alzheimer's disease, hypertension and muscular
CC dystrophy. The present sequence represents the DNA encoding the
CC SLAP1 fusion protein of the invention.
XX
SQ Sequence 1525 BP; 467 A; 286 C; 314 G; 458 T; 0 other;

Alignment Scores:
Pred. No.: 5.02e-07 Length: 1525
Score: 143.00 Matches: 27
Percent Similarity: 55.36% Conservative: 4
Best Local Similarity: 48.21% Mismatches: 23
Query Match: 43.73% Indels: 2
DB: 24 Gaps: 1

09-833799-13B (1-57) x ABK88022 (1-1525)

QY 3 GluProValLysGlyProValSerThr-----LysProGlySerCysProIleIleLeu 20
Db 156 GACCCAGTTGACACCCCAACCCAACTAGAGAAAGCCAGGTAAGTCCAGTTACTTAC 215
QY 21 IleArgCysAlaMetLeuAsnProProAsnArgCysLeuLysAspThrAspCysProGly 40
Db 216 GGTCAATGTTGATGTTGAACCCCACTTCTGTGAAATGGACGGTCAATGTAGAGA 275
QY 41 IleLysLysCysCysGluGlySerCysGlyMetAlaCysPheValPro 56
Db 276 GACTTGAAGTGTGATGGTATGTTGGTAACTCTGTTGTTCTGTTCTTCCCA 323

RESULT 50
ABK88025
ID ABK88025 standard; DNA; 1525 BP.
XX
AC ABK88025;
XX
DT 07-OCT-2002 (first entry)
XX

DE DNA sequence encoding rSLAP1 fusion protein.
XX
KW rSLAP1; gene; ds; Alzheimer's disease; tumour angiogenesis;
KW malaria; emphysema; asthma; chronic obstructive pulmonary disease;
KW cystic fibrosis; otitis media; otitis externa; HIV; psoriasis; eczema;
KW human immunodeficiency virus; atopic dermatitis; muscular dystrophy;
KW herpes; ulceration; sepsis; rheumatoid arthritis; periodontal disease;
KW tumour metastasis; osteoporosis; Paget's disease; scleroderma;
KW glomerulonephritis; hypertension.
XX
OS Homo sapiens.
OS Synthetic.
XX
FH Key Location/Qualifiers
FT RBS 6..8
FT /*tag= a
FT /standard_name= "Ribosome binding site"
FT CDS 9..1520
FT /*tag= b
FT /product= "rSLAP1 fusion protein"
FT misc_feature 12..1193
FT /*tag= c
FT /note= "AAT coding region"
FT misc_feature 1194..1196
FT /*tag= d
FT /note= "linking codon"
FT misc_feature 1197..1517
FT /*tag= e
FT /note= "SLPI coding region"
XX
WO200250287-A2.
XX
27-JUN-2002.
XX
18-DEC-2001; 2001WO-US49256.
XX
18-DEC-2000; 2000US-256699P.
XX 20-NOV-2001; 2001US-331966P.
XX (ARRI-) ARRIVA PHARM INC.
XX
PI Barr PJ, Gibson HL, Pemberton P;
XX
WPI; 2002-500631/53.
XX P-PSDB; AAU99884.
XX
Novel fusion protein useful for inhibiting protease activity associated
PT with a disorder such as emphysema, asthma, comprises a first protease
PT inhibitor comprising alpha 1-antitrypsin and a second protease
PT inhibitor -
XX
PS Example 3; Page 89-90; 134pp; English.
XX This invention relates to a novel fusion protein comprising a first
CC protease inhibitor comprising an alpha1-antitrypsin or its functionally
CC active portion and a second protease inhibitor or its functionally
CC active protein. The fusion proteins of the invention may act as an
CC inhibitor of protease activity. The fusion protein of the invention
CC is useful for inhibiting protease activity associated with a disorder
CC such as emphysema, asthma, chronic obstructive pulmonary disease,
CC cystic fibrosis, otitis media, otitis externa or HIV infection, or
CC for treating an individual suffering from or at risk for a disease or
CC disorder involving unwanted protease activity. The proteins are useful
CC for treating dermatological diseases such as atopic dermatitis, eczema
CC and psoriasis, in inflammatory responses to viral infection, and for
CC treating herpes infection, corneal or epidermal ulceration, chronic
CC non-healing wounds, sepsis, rheumatoid arthritis, periodontal disease,
CC tumour metastasis and tumour angiogenesis, gastric ulceration,
CC osteoporosis, Paget's disease, glomerulonephritis, scleroderma, malaria,
CC bacterial infection, Alzheimer's disease, hypertension and muscular
CC dystrophy. The present sequence represents the DNA encoding the
CC rSLAP1 fusion protein of the invention.
XX

sequences used in the isolation and characterisation of the proteins.
Primers AAC97535 - AAC97573 and AAC95582 - AAC97613 are used in the
construction of DNA encoding the protease inhibitors. Peptide sequences
AAB53098 - AAB53122 represent the protease inhibitor of the invention and
various other peptides used in the isolation of the protease inhibitors.
The serine protease inhibitors have cytostatic and anti-inflammatory activity.
The serine protease inhibitor protein is useful for treating a
protease-mediated condition, which includes protease mediated tissue
destruction, e.g. emphysema, arthritis, glomerulonephritis,
periodontitis, muscular dystrophy or tumour invasion. It is also useful
for modulating protease activity. The DNAs are useful for producing the
proteins, especially by recombinant methods.

Sequence 2274 BP; 459 A; 567 C; 602 G; 646 T; 0 other;

Alignment Scores:

Pred. No.: 1.39e-06 Length: 2274
Score: 141.00 Matches: 24
Percent Similarity: 55.10% Conservative: 3
Best Local Similarity: 48.98% Mismatches: 22
Query Match: 43.12% Indels: 0
DB: 21 Gaps: 0

09-833799-13B (1-57) x AAC97530 (1-2274)

QY 8 ProValSerThrLysProGlySerCysProIleLeuIleArgCysAlaMetLeuAsn 27
Db 1454 CCAGCAAGGAGGAGCTGGGAAGTGGCCAGTGACTTATGGCCAAATGTTGATGCTTAAC 1513
QY 28 ProProAsnArgCysLeuLysAspThrAspCysProGlyIleLysLysCysCysGluGly 47
Db 1514 CCCCCCAATTTCTGTGAGATGGATGGCCAGTGCACAGCGTGACTTGAAGTGTTCATGGGC 1573
QY 48 SerCysGlyMetAlaCysPheValPro 56
Db 1574 ATGTGTGGGAATCTCGCTTTCCTTCCCT 1600

RESULT 53

AAI67491
ID AAI67491 standard; DNA; 2274 BP.

AC AAI67491;

DT 11-FEB-2002 (first entry)

Recombinant secretory leukocyte protease inhibitor (SLPI) DNA.

Serine protease inhibitor protein; recombinant; leukocyte elastase;
trypsin; secretory leukocyte protease inhibitor; SLPI; ds.

Homo sapiens.

Key Location/Qualifiers
CDS 884..1607
/*tag= a
/product= "SLPI"
/note= "contains introns"
exon 884..1039
/*tag= b
intron 1040..1459
/*tag= c
exon 1460..1607
/*tag= d

US6291662-B1.

18-SEP-2001.

22-SEP-1998; 98US-0158085.

30-MAR-1987; 87US-0031846.

06-AUG-1990; 90US-0563832.

22-JUL-1994; 94US-0279056.

PR 05-DEC-1984; 84US-0678222.
PR 02-DEC-1985; 85US-0803471.
PR 29-JUL-1986; 86US-0890526.
XX
PA (AMGE-) AMGEN INC.
XX
PI Bandyopadhyay PK, Eisenberg SP, Stetler GL, Thompson RC;
XX
XX WPI; 2001-637974/73.
DR P-PSDB; AAG65996.
DR
XX
PT New DNA sequences, useful in recombinant DNA techniques for directing
PT the production of a serine protease inhibitor protein, e.g. leukocyte
PT elastase or trypsin
XX
PS Disclosure; Columns 13-18; 37pp; English.
XX
CC The invention relates to a DNA sequence encoding an analog of a mammalian
CC serine protease inhibitor protein. The analog comprises at least eight
CC cysteine residues and possesses serine protease inhibitor activity, and
CC at least one operational element consisting of a promoter, an operator,
CC a leader sequence, a Shine-Dalgarno sequence, a ribosome binding site
CC or a terminator codon. The DNA is useful in recombinant DNA techniques
CC for directing the production of a serine protease inhibitor protein,
CC e.g. leukocyte elastase or trypsin. The present sequence represents the
CC DNA that encodes a secretory leukocyte protease inhibitor (SLPI).
XX
SQ Sequence 2274 BP; 459 A; 567 C; 602 G; 646 T; 0 other;

Alignment Scores:

Pred. No.: 1.39e-06 Length: 2274
Score: 141.00 Matches: 24
Percent Similarity: 55.10% Conservative: 3
Best Local Similarity: 48.98% Mismatches: 22
Query Match: 43.12% Indels: 0
DB: 22 Gaps: 0

09-833799-13B (1-57) x AAI67491 (1-2274)

QY 8 ProValSerThrLysProGlySerCysProIleLeuIleArgCysAlaMetLeuAsn 27
Db 1454 CCAGCAAGGAGGAGCTGGGAAGTGGCCAGTGACTTATGGCCAAATGTTGATGCTTAAC 1513
QY 28 ProProAsnArgCysLeuLysAspThrAspCysProGlyIleLysLysCysCysGluGly 47
Db 1514 CCCCCCAATTTCTGTGAGATGGATGGCCAGTGCACAGCGTGACTTGAAGTGTTCATGGGC 1573
QY 48 SerCysGlyMetAlaCysPheValPro 56
Db 1574 ATGTGTGGGAATCTCGCTTTCCTTCCCT 1600

RESULT 54

AAQ45442
ID AAQ45442 standard; DNA; 180 BP.
XX
AC AAQ45442;
XX
DT 26-OCT-1994 (first entry)
XX
DE CLPI.

Serine leukocyte protease inhibitor; SLPI; truncated SLPI; CLPI;
retrovirus; retroviral infection; HIV; parotid inhibitor; enzyme;
elastase; cathepsin G; trypsin; ss.

Homo sapiens.

Key Location/Qualifiers
misc_feature 79..81
/*tag= a
/transl_except= pos:79..81, aa:Leu

PN W09406454-A.

XX 31-MAR-1994.
XX 09-SEP-1993; 93WO-US08486.
XX 09-SEP-1992; 92US-0943369.
XX (SYND) SYNERGEN INC.
XX (USSH) US DEPT HEALTH & HUMAN SERVICES.

Eisenberg S, Thompson RC, Wahl SM;

WPI; 1994-118153/14.
P-PSDB; AAR51172.

Inhibiting retroviral infection with serine leukocyte protease inhibitor - esp. for treating or preventing HIV infection, also new protease inhibitors and nucleic acids encoding them

Claim 29; Page 17; 21pp; English.

Retroviral infection is inhibited by admin. of an SLPI, or its analogue or deriv. A useful SLPI deriv. is CLPI, a truncated SLPI mol. having only the last 60 amino acids of the native parotid inhibitor. SLPI acts by blocking a host cell enzyme essential for retroviral infection; it is a powerful inhibitor of human leukocyte elastase, cathepsin G and human trypsin. The inhibitor provides an effective and relatively safe method for treating HIV infections.

Sequence 180 BP; 42 A; 45 C; 50 G; 43 T; 0 other;

Alignment Scores:
Pred. No.: 1.06e-07 Length: 180
Score: 139.00 Matches: 26
Percent Similarity: 53.57% Conservative: 4
Best Local Similarity: 46.43% Mismatches: 24
Query Match: 42.51% Indels: 2
DB: 15 Gaps: 1

09-833799-13B (1-57) x AAQ45442 (1-180)

QY 3 GluProValLysGlyProValSerThr-----LysProGlySerCysProIleLeu 20
Db 4 GATCCTGTTGACACCCCAACACCAAGGAGGAGCTGGGAAGTGCCAGTGACTTAT 63
QY 21 IleArgCysAlaMetLeuAsnProProAsnArgCysLeuLysAspThrAspCysProGly 40
Db 64 GGCCAATGTTTGATGCCCTAACCCCCCAATTTCTGTGAGATGGATGGCCAGTGCAAGCGT 123
QY 41 IleLysLysCysCysGluGlySerCysGlyMetAlaCysPheValPro 56
Db 124 GACTTGAAGTGTTCATGGGCATGTGTGGGAAATCCTGCGTTTCCCT 171

RESULT 55
AAZ88483
ID AAZ88483 standard; DNA; 180 BP.
XX
AC AAZ88483;
XX
DT 12-MAY-2000 (first entry)
XX
DE Truncated SLPI encoding nucleotide sequence CLPI SEQ ID NO:2.

Secretory leukocyte protease inhibitor; retroviral; infection; HIV; AIDS; SLPI; CLPI; parotid inhibitor; antiviral; anti-HIV; ss.

Homo sapiens.

US6017880-A.

25-JAN-2000.

07-JUN-1995; 95US-0483503.

XX 09-SEP-1992; 92US-0943369.
XX 09-SEP-1993; 93WO-US08486.
XX 09-MAR-1994; 94US-0209040.
XX (AMGE-) AMGEN INC.
XX (USSH) US DEPT HEALTH & HUMAN SERVICES.
XX
PI Dripps DJ, Thompson RC, Wahl SM, Eisenberg S;
XX
DR WPI; 2000-170378/15.
DR P-PSDB; AAY79951.

Inhibiting retrovirus infection using secretory leukocyte protease inhibitors for treating diseases such as acquired immunodeficiency syndrome -

Disclosure; Column 3-5; 18pp; English.

The present invention describes a method for inhibiting, in vitro, retrovirus infection of CD4+ cells comprising contacting the cells with a secretory leukocyte protease inhibitor (SLPI) having the amino acid sequence of a naturally occurring SLPI or a substitution analogue having 107 amino acid residues. SLPI inhibits enzymes whose function are necessary for retroviral infection of the cell. SLPI is a potent inhibitor of elastase, trypsin, cathepsin G and chymotrypsin. The method is useful for treating retrovirus infection in a patient and associated diseases such as acquired immunodeficiency syndrome (AIDS). The present sequence encodes a truncated version of SLPI designated CLPI.

Sequence 180 BP; 42 A; 45 C; 50 G; 43 T; 0 other;

Alignment Scores:
Pred. No.: 1.06e-07 Length: 180
Score: 139.00 Matches: 26
Percent Similarity: 53.57% Conservative: 4
Best Local Similarity: 46.43% Mismatches: 24
Query Match: 42.51% Indels: 2
DB: 21 Gaps: 1

09-833799-13B (1-57) x AAZ88483 (1-180)

QY 3 GluProValLysGlyProValSerThr-----LysProGlySerCysProIleLeu 20
Db 4 GATCCTGTTGACACCCCAACACCAAGGAGGAGCTGGGAAGTGCCAGTGACTTAT 63
QY 21 IleArgCysAlaMetLeuAsnProProAsnArgCysLeuLysAspThrAspCysProGly 40
Db 64 GGCCAATGTTTGATGCCCTAACCCCCCAATTTCTGTGAGATGGATGGCCAGTGCAAGCGT 123
QY 41 IleLysLysCysCysGluGlySerCysGlyMetAlaCysPheValPro 56
Db 124 GACTTGAAGTGTTCATGGGCATGTGTGGGAAATCCTGCGTTTCCCT 171

RESULT 56
AAZ88483
ID AAZ88483 standard; DNA; 1010 BP.
XX
AC AAZ88483;
XX
DT 24-OCT-2000 (first entry)
XX
DE Human peptidase, HPEP-9 coding sequence.

Human; peptidase; cell proliferative disorder; arteriosclerosis; psoriasis; myelofibrosis; cancer; autoimmune disorder; Crohn's disease; inflammatory disorder; AIDS; anaemia; allergy; asthma; atherosclerosis; Grave's disease; multiple sclerosis; scleroderma; infection; diabetes; metabolic disorder; Addison's disease; cystic fibrosis; diagnosis; glycogen storage disease; obesity; therapy; HPEP-9; ds.

Homo sapiens.

XX

```
FH Key Location/Qualifiers
FT CDS 90..461
FT /*tag= a
FT /product= HPEP-9
XX
PN WO200042201-A2.
XX
XX 20-JUL-2000.
XX
XX 11-JAN-2000; 2000WO-US00641.
XX
PR 11-JAN-1999; 99US-0172247.
PR 03-MAY-1999; 99US-0132253.
PR 27-MAY-1999; 99US-0136653.
XX
XX (INCY-) INCYTE PHARM INC.
XX
XX Bandman O, Hillman JL, Tang YT, Azimzai Y, Baughn MR, Lal P;
PI Yue H, Lu DAM;
XX
DR WPI; 2000-482832/42.
DR P-PSDB; AAY90292.
XX
XX An isolated polypeptide for diagnosis, prevention and treatment of
PT cell proliferative, autoimmune/ inflammatory and metabolic disorders
PT comprises a sequence encoding a human peptidase .
XX
PS Claim 4; Page 118; 131pp; English.
XX
CC This sequence encodes the human peptidase, designated HPEP-9. The
CC invention relates to 18 human peptidases designated HPEP-1 to HPEP-18,
CC respectively. The peptidases can be used for treating a disease or
CC condition associated with decreased expression or over expression of
CC functional human peptidases. The diseases that can be diagnosed,
CC prevented and treated include cell proliferative disorders (such as
CC arteriosclerosis, psoriasis, myelofibrosis, and cancers),
CC autoimmune/inflammatory disorders (such as AIDS, anaemia, allergies,
CC Crohn's disease, asthma, atherosclerosis, Grave's disease, multiple
CC sclerosis, and scleroderma), infections, and metabolic disorders (such as
CC Addison's disease, diabetes, cystic fibrosis, glycogen storage diseases
CC and obesity).
XX
SQ Sequence 1010 BP; 223 A; 308 C; 243 G; 236 T; 0 other;

Alignment Scores:
Pred. No.: 5.4e-05 Length: 1010
Score: 123.50 Matches: 25
Percent Similarity: 61.22% Conservative: 5
Best Local Similarity: 51.02% Mismatches: 16
Query Match: 37.77% Indels: 3
DB: 21 Gaps: 2

09-833799-13B (1-57) x AAA37665 (1-1010)
QY 9 ValSerThrLysProGlySerCysProIleLeuIleArgCysAlaMetLeuAsnPro 28
Db 309 GTCTCTGTGAAGCTGGGCGAGCTGCCAGAGGACCAACTGGCTGC-----CTCAGCCCC 362
QY 29 ProAsnArg---CysLeuLysAspThrAspCysProGlyIleLysCysCysGluGly 47
Db 363 ATGAACCACTGTGTTACAAGGACTCAGACTGCTCGGGCAAAAGCGATGCTGCCACAGC 422
QY 48 SerCysGlyMetAlaCysPheValPro 56
Db 423 GCCTGGGGCGGGATTGCCGGGATCCT 449

RESULT 57
AAD38698
ID AAD38698 standard; cDNA; 1013 BP.
XX
AC AAD38698;
XX
DT 23-SEP-2002 (first entry)
```

```
XX Human LP229 secreted protein encoding cDNA.
DE
XX
KW Human; secreted protein; atherosclerosis; Alzheimer's disease; LP229;
KW diabetic retinopathy; severe combined immunodeficiency; pancreatitis;
KW rheumatoid arthritis; colorectal adenoma; haemolytic anaemia; cancer;
KW reperfusion injury; arteriosclerosis; wound healing; transgenic animal;
KW gene therapy; neoplasm; transgenic; psoriasis; ischaemia; carcinoma;
KW chromosome 20q12-13.2; gene; ss.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT CDS 90..461
FT /*tag= a
FT /product= "Human LP229 secreted protein"
FT sig_peptide 90..143
FT /*tag= b
FT mat_peptide 144..458
FT /*tag= c
FT /product= "Mature human LP229 secreted protein"
XX
XX WO200226801-A2.
XX
XX 04-APR-2002.
XX
XX 14-SEP-2001; 2001WO-US26026.
XX
XX 28-SEP-2000; 2000US-236088P.
XX (ELIL ) LILLY & CO ELI.
XX
XX Su EW, Wang H;
XX
XX WPI; 2002-471259/50.
XX P-PSDB; AAE23983.
XX
XX Novel proteins and polynucleotides of secreted proteins useful for
XX treating various diseases e.g. rheumatoid arthritis, cancer, psoriasis,
XX diabetic retinopathy, arteriosclerosis, ischemia or reperfusion injury
XX
PS Claim 1; Page 138-139; 145pp; English.
XX
CC The invention relates to human secreted polypeptides designated LP095,
CC LP191, LP217, LP220, LP221, LP222, LP229, LP237 or LP238 and nucleic
CC acid molecules encoding such polypeptides. Novel secreted proteins of
CC the invention are used for treating diseases such as atherosclerosis,
CC Alzheimer's disease, diabetic retinopathy, psoriasis, pancreatitis,
CC arteriosclerosis, rheumatoid arthritis, colorectal adenomas, severe
CC combined immunodeficiency, ischaemia, carcinomas, haemolytic anaemia,
CC reperfusion injury, neoplasms and cancer especially liver cancer. They
CC are also used for wound healing. Polynucleotides of the invention can
CC be used to generate transgenic animals or knock out animals, which in
CC turn, are useful in the development and screening of therapeutically
CC useful reagents for use in the treatment of diseases associated with
CC LP polypeptide associated activity. They are also used in gene therapy.
CC The present sequence is human LP229 secreted protein encoding cDNA.
CC LP229 gene is located on chromosome 20q12-13.2.
XX
SQ Sequence 1013 BP; 226 A; 308 C; 243 G; 236 T; 0 other;

Alignment Scores:
Pred. No.: 5.42e-05 Length: 1013
Score: 123.50 Matches: 25
Percent Similarity: 61.22% Conservative: 5
Best Local Similarity: 51.02% Mismatches: 16
Query Match: 37.77% Indels: 3
DB: 24 Gaps: 2

09-833799-13B (1-57) x AAD38698 (1-1013)
QY 9 ValSerThrLysProGlySerCysProIleLeuIleArgCysAlaMetLeuAsnPro 28
```

Db 309 GTCTCTGTGAAGCTGGCAGCTGCCAGAGGACCAACTGCGTGC-----CTCAGCCCC 362
QY 29 ProAsnArg---CysLeuLysAspThrAspCysProGlyIleLysLysCysCysGluGly 47
Db 363 ATGAACCACTGTGTACAGGACTCAGACTGCTCGGGCAAAAGCGATGCTGCCACAGC 422
QY 48 SerCysGlyMetAlaCysPheValPro 56
Db 423 GCCTGGGGCGGGATTGCCGGGATCCT 449
RESULT 58
AAC97529
ID AAC97529 standard; DNA; 321 BP.
XX
AC AAC97529;
DT 27-FEB-2001 (first entry)
XX
DE DNA encoding a protease inhibitory peptide.
XX
KW Serine protease inhibitor; cytostatic; anti-inflammatory; arthritis;
KW protease mediated tissue destruction; emphysema; glomerulonephritis;
KW periodontitis; muscular dystrophy; tumour invasion; chymotrypsin;
KW elastase; ds.
XX
OS Synthetic.
XX
PN US6132990-A.
XX
PD 17-OCT-2000.
XX
PF 07-JUN-1991; 91US-0712354.
XX
PR 03-JAN-1989; 89US-0293042.
PR 06-DEC-1984; 84US-0678822.
PR 02-DEC-1985; 85US-0803471.
PR 29-JUL-1986; 86US-0890526.
PR 30-MAR-1987; 87US-0031846.
PR 04-AUG-1987; 87US-0082962.
XX
PA (AMGE-) AMGEN BOULDER INC.
XX
PI Bandyopadhyay PK, Eisenberg SP, Stetler GL, Thompson RC;
XX WPI; 2000-678667/66.
XX
PT New serine protease inhibitors and DNA sequences for treating a
PT protease-mediated condition or tissue destruction e.g. emphysema or
PT tumor invasion and for recombinant production of inhibitors -
XX
PS Disclosure; Column 12-13; 47pp; English.
XX
CC This invention relates to new purified and isolated mammalian serine
CC protease inhibitor proteins which comprise at least 8 cysteine residues
CC and no more than 107 amino acids. The protease inhibitors are capable of
CC inhibiting chymotrypsin and elastase. Sequences AAC97526 - AAC97534,
CC AAC97574 - AAC97581 and AAC97614 - AAC97641 represent DNA encoding the
CC protease inhibitors of the invention, and include oligonucleotide
CC sequences used in the isolation and characterisation of the proteins.
CC Primers AAC97535 - AAC97573 and AAC95582 - AAC97613 are used in the
CC construction of DNA encoding the protease inhibitors. Peptide sequences
CC AAB53098 - AAB53122 represent the protease inhibitor of the invention and
CC various other peptides used in the isolation of the protease inhibitors.
CC The protease inhibitors have cytostatic and anti-inflammatory activity.
CC The serine protease inhibitor protein is useful for treating a
CC protease-mediated condition, which includes protease mediated tissue
CC destruction, e.g. emphysema, arthritis, glomerulonephritis,
CC periodontitis, muscular dystrophy or tumour invasion. It is also useful
CC for modulating protease activity. The DNAs are useful for producing the
CC proteins, especially by recombinant methods.
XX
SQ Sequence 321 BP; 60 A; 54 C; 60 G; 44 T; 103 other;

Alignment Scores: 1.98e-05 Length: 321
Pred. No.: 122.00 Matches: 25
Score: 51.79% Conservative: 4
Percent Similarity: 44.64% Mismatches: 25
Best Local Similarity: 37.31% Indels: 2
Query Match: 21 Gaps: 1
DB: 21
09-833799-13B (1-57) x AAC97529 (1-321)
QY 3 GluProValLysGlyProValSerThr-----LysProGlySerCysProIleLeu 20
Db 145 GAYCNGTNGAYACNCCNAAAYCCNACNCGNCGNAAARTGNCNGTNACTAY 204
QY 21 IleArgCysAlaMetLeuAsnProProAsnArgCysLeuLysAspThrAspCysProGly 40
Db 205 GGNCARTGYCTNATGCTNAAAYCCNCCNAAAYTTTGTGNGARATGGAYGGNCARTGYAARCGN 264
QY 41 IleLysLysCysCysGluGlySerCysGlyMetAlaCysPheValPro 56
Db 265 GAYCTNAAARTGYTGATGGGNATNTGYGGNAAARTCNTGTGTTNTCNCN 312
RESULT 59
AAZ51673
ID AAZ51673 standard; cDNA; 1411 BP.
XX
AC AAZ51673;
XX
DT 21-JUN-2000 (first entry)
XX
DE Human p53 target molecule, PRG5 cDNA.
XX
KW PRG5; p53 target; human; modulate; cell proliferation; immunomodulatory;
KW chromosome 20q12-13; cytostatic; gene therapy; tumour cell; inducer;
KW diagnosis; therapeutic; proliferative disease; cell cycle arrest; cancer;
KW treatment; apoptosis; knockout animal; cancer susceptibility; ALP;
KW antileukoproteinase; secretory leukocyte proteinase inhibitor; SLPI; ss.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT 5'UTR 1..64
FT /*tag= a
FT CDS 65..436
FT /*tag= b
FT /product= "Human PRG5 protein"
FT /function= "Proteinase inhibitor activity"
FT sig_peptide 65..112
FT /*tag= b
FT mat_peptide 113..433
FT /*tag= c
FT /product= "Mature human PRG5 protein"
FT 3'UTR 437..1411
FT /*tag= c
XX WO200012526-A1.
XX 09-MAR-2000.
XX 27-AUG-1999; 99WO-US19551.
XX 28-AUG-1998; 98US-0098251.
XX (UYPR-) UNIV PRINCETON.
XX Horikoshi N, Shenk T;
XX WPI; 2000-246724/21.
XX P-PSDB; AAY70471.
XX New p53-inducible isolated nucleic acid molecule including open reading

XX 26-NOV-2001 (first entry)
XX Nucleotide sequence of a human secreted polypeptide.
DE Human; secreted polypeptide; nervous disease; muscular disease; tumour;
XX gastrointestinal ulceration; spinal cord disease; trachea disease;
KW thyroid gland disease; ovary disease; prostate disease; heart disease;
KW renal gland disease; small intestine disease; thymus disease;
KW lymph node disease; muscular system disease; colon disease;
KW lipase deficiency; cystic fibrosis; pancreatitis; clot formation;
KW myocardial infarction; angioplasty; liver disease; coagulation disorder;
KW microbial disease; immune disorder; inflammation; transplant rejection;
KW bone thickness; bone density; ferroxidase loss; apoptosis;
KW vascular smooth cell proliferation; vaccine; ss.
XX
OS Homo sapiens.
XX
XX Key Location/Qualifiers
FH CDS 1..732
FT /*tag= a
FT /product= "secreted polypeptide"
XX
XX WO200166690-A2.
XX 13-SEP-2001.
XX
XX 05-MAR-2001; 2001WO-US07143.
XX
XX 06-MAR-2000; 2000US-0187107.
PR 13-MAR-2000; 2000US-0188916.
PR 03-OCT-2000; 2000US-0236874.
PR 03-OCT-2000; 2000US-0237846.
XX
XX (SMIK) SMITHKLINE BEECHAM CORP.
PA (SMIK) SMITHKLINE BEECHAM PLC.
PA
XX
PI Agarwal P, Murdoch PR, Rizvi SK, Smith RF, Xiang Z;
XX
DR WPI; 2001-570768/64.
DR P-PSDB; AAG67518.
XX
PT Novel isolated secreted polypeptide useful for treating nervous and
PT muscular diseases, gastrointestinal ulceration, coagulation and immune
PT disorders, microbial diseases, inflammation and transplant rejection -
XX
PS Claim 2; Page 47; 102pp; English.
XX
CC The present sequence encodes a human secreted polypeptide. The
CC secreted polypeptides and polynucleotides are useful for treating
CC nervous and muscular diseases, for inhibiting tumour formation and
CC metastasis, for treating gastrointestinal ulceration, for preventing
CC and treating diseases in spinal cord, thyroid gland, ovary, prostate,
CC renal gland, small intestine, heart, trachea, thymus, lymph node,
CC muscular system and colon, for treating lipase deficiency in cystic
CC fibrosis and pancreatitis, for treating undesirable clot formation
CC such as myocardial infarction, during angioplasty and all surgical
CC procedures that require decreased blood clot formation, for treating
CC liver diseases, coagulation disorders and microbial diseases, for
CC treating immune disorders, for treating inflammation and transplant
CC rejection, for enhancing bone thickness and increasing bone density,
CC for reducing the loss of essential ferroxidases, for suppressing
CC apoptosis, and for regulating vascular smooth cell proliferation. They
CC may also be used as vaccines.
XX
SQ Sequence 732 BP; 149 A; 212 C; 215 G; 156 T; 0 other;

Alignment Scores:
Pred. No.: 0.000138 Length: 732
Score: 118.50 Matches: 24
Percent Similarity: 61.22% Conservative: 6
Best Local Similarity: 48.98% Mismatches: 16
Query Match: 36.24% Indels: 3

DB: 22 Gaps: 2
09-833799-13B (1-57) x AAH78210 (1-732)
Qy 9 ValSerThrLysProGlySerCysProIleIleLeuIleArgCysAlaMetLeuAsnPro 28
Db 530 GTTCAGTGAAGCTGGGAGCTGCCAGAGGACCAACTGCGCTGC-----CTCAGCCCC 583
Qy 29 ProAsnArg---CysLeuLysAspThrAspCysProGlyIleLysLysCysCysGluGly 47
Db 584 ATGAACCACTGTGTCAACAAGGACTCAGACTGCTCGGGCAAAAGCGATGCTGCCACAGC 643
Qy 48 SerCysGlyMetAlaCysPheValPro 56
Db 644 GCCTGCGGGCGGATTGCGGGGATCCT 670
RESULT 62
AAL07251
ID AAL07251 standard; DNA; 5690 BP.
XX
AC AAL07251;
XX
DT 21-NOV-2001 (first entry)
XX
DE Human reproductive system related antigen DNA SEQ ID NO: 9939.
XX
KW Human; reproductive system related antigen; reproductive system disorder;
KW cancer; gene therapy; ds.
XX
OS Homo sapiens.
XX
PN WO200155320-A2.
XX
PD 02-AUG-2001.
XX
PF 17-JAN-2001; 2001WO-US01339.
XX
PR 31-JAN-2000; 2000US-0179065.
PR 04-FEB-2000; 2000US-0180628.
PR 24-FEB-2000; 2000US-0184664.
PR 02-MAR-2000; 2000US-0186350.
PR 16-MAR-2000; 2000US-0189874.
PR 17-MAR-2000; 2000US-0190076.
PR 18-APR-2000; 2000US-0198123.
PR 19-MAY-2000; 2000US-0205515.
PR 07-JUN-2000; 2000US-0209467.
PR 28-JUN-2000; 2000US-0214886.
PR 30-JUN-2000; 2000US-0215135.
PR 07-JUL-2000; 2000US-0216647.
PR 07-JUL-2000; 2000US-0216880.
PR 11-JUL-2000; 2000US-0217487.
PR 11-JUL-2000; 2000US-0217496.
PR 14-JUL-2000; 2000US-0218290.
PR 26-JUL-2000; 2000US-0220963.
PR 26-JUL-2000; 2000US-0220964.
PR 14-AUG-2000; 2000US-0224518.
PR 14-AUG-2000; 2000US-0224519.
PR 14-AUG-2000; 2000US-0225213.
PR 14-AUG-2000; 2000US-0225214.
PR 14-AUG-2000; 2000US-0225266.
PR 14-AUG-2000; 2000US-0225267.
PR 14-AUG-2000; 2000US-0225268.
PR 14-AUG-2000; 2000US-0225270.
PR 14-AUG-2000; 2000US-0225447.
PR 14-AUG-2000; 2000US-0225757.
PR 14-AUG-2000; 2000US-0225758.
PR 14-AUG-2000; 2000US-0225759.
PR 18-AUG-2000; 2000US-0226279.
PR 22-AUG-2000; 2000US-0226681.
PR 22-AUG-2000; 2000US-0226868.
PR 22-AUG-2000; 2000US-0227182.
PR 23-AUG-2000; 2000US-0227009.
PR 30-AUG-2000; 2000US-0228924.

AC ABL98799;
XX 21-JUN-2002 (first entry)
DT Human testicular antigen encoding DNA fragment SEQ ID NO: 3451.
XX
DE Human; testicular antigen; testes; cancer; metastasis; immune disorder;
KW reproductive system disorder; urinary system disorder; gene therapy;
KW cardiovascular disorder; respiratory disorder; neurological disorder;
KW gastrointestinal disease; infection; cytostatic; gene; ds.
XX
OS Homo sapiens.
XX WO200155317-A2.
PN 02-AUG-2001.
XX
PD 17-JAN-2001; 2001WO-US01329.
XX
PR 31-JAN-2000; 2000US-0179065.
PR 04-FEB-2000; 2000US-0180628.
PR 24-FEB-2000; 2000US-0184664.
PR 02-MAR-2000; 2000US-0186350.
PR 16-MAR-2000; 2000US-0189874.
PR 17-MAR-2000; 2000US-0190076.
PR 18-APR-2000; 2000US-0198123.
PR 19-MAY-2000; 2000US-0205515.
PR 07-JUN-2000; 2000US-0209467.
PR 28-JUN-2000; 2000US-0214886.
PR 30-JUN-2000; 2000US-0215135.
PR 07-JUL-2000; 2000US-0216647.
PR 07-JUL-2000; 2000US-0216880.
PR 11-JUL-2000; 2000US-0217487.
PR 11-JUL-2000; 2000US-0217496.
PR 14-JUL-2000; 2000US-0218290.
PR 26-JUL-2000; 2000US-0220963.
PR 26-JUL-2000; 2000US-0220964.
PR 14-AUG-2000; 2000US-0224518.
PR 14-AUG-2000; 2000US-0224519.
PR 14-AUG-2000; 2000US-0225213.
PR 14-AUG-2000; 2000US-0225214.
PR 14-AUG-2000; 2000US-0225266.
PR 14-AUG-2000; 2000US-0225267.
PR 14-AUG-2000; 2000US-0225268.
PR 14-AUG-2000; 2000US-0225270.
PR 14-AUG-2000; 2000US-0225447.
PR 14-AUG-2000; 2000US-0225757.
PR 14-AUG-2000; 2000US-0225758.
PR 14-AUG-2000; 2000US-0225759.
PR 18-AUG-2000; 2000US-0226279.
PR 22-AUG-2000; 2000US-0226681.
PR 22-AUG-2000; 2000US-0226868.
PR 22-AUG-2000; 2000US-0227182.
PR 23-AUG-2000; 2000US-0227009.
PR 30-AUG-2000; 2000US-0228924.
PR 01-SEP-2000; 2000US-0229287.
PR 01-SEP-2000; 2000US-0229343.
PR 01-SEP-2000; 2000US-0229344.
PR 01-SEP-2000; 2000US-0229345.
PR 05-SEP-2000; 2000US-0229509.
PR 06-SEP-2000; 2000US-0229513.
PR 06-SEP-2000; 2000US-0230437.
PR 06-SEP-2000; 2000US-0230438.
PR 08-SEP-2000; 2000US-0231242.
PR 08-SEP-2000; 2000US-0231243.
PR 08-SEP-2000; 2000US-0231244.
PR 08-SEP-2000; 2000US-0231413.
PR 08-SEP-2000; 2000US-0231414.
PR 08-SEP-2000; 2000US-0232080.
PR 08-SEP-2000; 2000US-0232081.
PR 12-SEP-2000; 2000US-0231968.
PR 14-SEP-2000; 2000US-0232397.
PR 14-SEP-2000; 2000US-0232398.
PR 14-SEP-2000; 2000US-0232399.
PR 14-SEP-2000; 2000US-0232400.
PR 14-SEP-2000; 2000US-0232401.
PR 14-SEP-2000; 2000US-0233063.
PR 14-SEP-2000; 2000US-0233064.
PR 14-SEP-2000; 2000US-0233065.
PR 21-SEP-2000; 2000US-0234223.
PR 21-SEP-2000; 2000US-0234274.
PR 25-SEP-2000; 2000US-0234997.
PR 25-SEP-2000; 2000US-0234998.
PR 26-SEP-2000; 2000US-0235484.
PR 27-SEP-2000; 2000US-0235834.
PR 27-SEP-2000; 2000US-0235836.
PR 29-SEP-2000; 2000US-0236327.
PR 29-SEP-2000; 2000US-0236367.
PR 29-SEP-2000; 2000US-0236368.
PR 29-SEP-2000; 2000US-0236369.
PR 29-SEP-2000; 2000US-0236370.
PR 02-OCT-2000; 2000US-0236802.
PR 02-OCT-2000; 2000US-0237037.
PR 02-OCT-2000; 2000US-0237038.
PR 02-OCT-2000; 2000US-0237039.
PR 02-OCT-2000; 2000US-0237040.
PR 13-OCT-2000; 2000US-0239935.
PR 13-OCT-2000; 2000US-0239937.
PR 20-OCT-2000; 2000US-0240960.
PR 20-OCT-2000; 2000US-0241221.
PR 20-OCT-2000; 2000US-0241785.
PR 20-OCT-2000; 2000US-0241786.
PR 20-OCT-2000; 2000US-0241787.
PR 20-OCT-2000; 2000US-0241808.
PR 20-OCT-2000; 2000US-0241809.
PR 20-OCT-2000; 2000US-0241826.
PR 01-NOV-2000; 2000US-0244617.
PR 08-NOV-2000; 2000US-0246474.
PR 08-NOV-2000; 2000US-0246475.
PR 08-NOV-2000; 2000US-0246476.
PR 08-NOV-2000; 2000US-0246477.
PR 08-NOV-2000; 2000US-0246478.
PR 08-NOV-2000; 2000US-0246523.
PR 08-NOV-2000; 2000US-0246524.
PR 08-NOV-2000; 2000US-0246525.
PR 08-NOV-2000; 2000US-0246526.
PR 08-NOV-2000; 2000US-0246527.
PR 08-NOV-2000; 2000US-0246528.
PR 08-NOV-2000; 2000US-0246532.
PR 08-NOV-2000; 2000US-0246609.
PR 08-NOV-2000; 2000US-0246610.
PR 08-NOV-2000; 2000US-0246611.
PR 08-NOV-2000; 2000US-0246613.
PR 17-NOV-2000; 2000US-0249207.
PR 17-NOV-2000; 2000US-0249208.
PR 17-NOV-2000; 2000US-0249209.
PR 17-NOV-2000; 2000US-0249210.
PR 17-NOV-2000; 2000US-0249211.
PR 17-NOV-2000; 2000US-0249212.
PR 17-NOV-2000; 2000US-0249213.
PR 17-NOV-2000; 2000US-0249214.
PR 17-NOV-2000; 2000US-0249215.
PR 17-NOV-2000; 2000US-0249216.
PR 17-NOV-2000; 2000US-0249217.
PR 17-NOV-2000; 2000US-0249218.
PR 17-NOV-2000; 2000US-0249244.
PR 17-NOV-2000; 2000US-0249245.
PR 17-NOV-2000; 2000US-0249264.
PR 17-NOV-2000; 2000US-0249265.
PR 17-NOV-2000; 2000US-0249297.
PR 17-NOV-2000; 2000US-0249299.
PR 17-NOV-2000; 2000US-0249300.
PR 01-DEC-2000; 2000US-0250160.
PR 01-DEC-2000; 2000US-0250391.
PR 05-DEC-2000; 2000US-0251030.
PR 05-DEC-2000; 2000US-0251988.

PF 04-DEC-1985; 85WO-US02385.
XX
PR 04-DEC-1985; 85WO-US02385.
PR 06-DEC-1984; 84US-0678822.
PR 02-DEC-1985; 85US-0803471.
XX
PA (SYNE-) SYNERGEN BIOLOG INC.
XX
PI Bandyopadhy PK, Eisenberg SP, Stetler GL, Thompson RC;
XX
DR WPI; 1986-169458/26.
XX
DR New synthetic DNA sequences for directing microbial synthesis -
PT for prodn. of single poly:peptide chain serine protease inhibitor
PT having leukocyte elastase and trypsin inhibitory sites
XX
PS Disclosure; Page 46; 59pp; English.
XX
CC The sequence directs synthesis of a secretory leukocyte protease-
CC inhibitor. See also AAN60463-67, AAN60469, and AAP60562-66.
XX
SQ Sequence 920 BP; 221 A; 239 C; 238 G; 221 T; 1 other;

Alignment Scores:
Pred. No.: 0.000354 Length: 920
Score: 116.00 Matches: 22
Percent Similarity: 40.30% Conservative: 5
Best Local Similarity: 32.84% Mismatches: 20
Query Match: 35.47% Indels: 20
DB: 7 Gaps: 1

09-833799-13B (1-57) x AAN60468 (1-920)
QY 10 SerThrLysProGlySerCysProIleLeuLeuLeuArgCysAlaMetLeuAsnProPro 29
Db 596 GCAGCAAAACCGGTAAATGCCCGGTAACTATGCCAGTGTCTGATGCTGAACCCGCG 655
QY 30 Asn----- 30
Db 656 AACTTTTGGCCCATTTACGGGCCATTGGATACCGGTACAGACTACCACTTGGCGGC 715
QY 31 ---ArgCysLeuLysAspThrAspCysProGlyLeuLysLysCysCysGluGlySerCys 49
Db 716 TTGAACGCGAAATGGACGGCCAGGTAAACACGACATCTGAAATGCTGTATGGGTATGTC 775
QY 50 GlyMetAlaCysPheValPro 56
Db 776 GGCAAGACGCTTTACTGCGG 796

RESULT 66
AAX80909
ID AAX80909 standard; cDNA; 478 BP.
XX
AC AAX80909;
XX
DT 03-NOV-1999 (first entry)
XX
DE Human cDNA clone HKABR62.
XX
KW Human cDNA clone HKABR62; secreted protein; serine protease inhibitor;
KW serine protease; Mus musculus secretory leukocyte protease inhibitor;
KW extracellular matrix degradation; multiple sclerosis; cancer; arthritis;
KW inflammation; immune system disorder; neurodegenerative disorder; serpin;
KW Kallmann's syndrome; Down's syndrome; Alzheimer's; fusion protein;
KW Galactorrhea; hypogonadism; somatostatin; protein purification; ss.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT CDS 19..249
FT /tag= a
FT /product= "Serine protease inhibitor"
FT sig_peptide 19..75

mat_peptide /*tag= b
76..246
/*tag= c
/label= Mature_serine_protease_inhibitor

WO9940183-A1.
12-AUG-1999.
04-FEB-1999; 99WO-US02292.
06-FEB-1998; 98US-0073961.
(HUMA-) HUMAN GENOME SCI INC.
Ni J, Ruben SM;
WPI; 1999-508502/42.
P-PSDB; AAY28645.

New isolated human serine protease and serpin polypeptides, used to develop products for treating e.g. immune disorders, cancers, inflammation, transplant rejection or infections, or as food additives

Claim 4; Page 88; 99pp; English.

The present sequence is a human cDNA clone HKABR62 which encodes serine protease inhibitor (serpin) and is obtained from human keratinocyte tissue cDNA library. The protein obtained from this clone shows a high degree of sequence similarity to Mus musculus secretory leukocyte protease inhibitor. The cDNA and the serpin are used in the diagnosis and treatment of disorders related to abnormal level of the protein or mutation in the nucleotide sequence. The serpin can be used for treating disorders characterised by degradation of extracellular matrix, e.g. cancer, arthritis, multiple sclerosis and immune system disorders, for treating wasting associated with excessive protease production during inflammation or neurodegenerative disorders e.g. Kallmann's and Down's syndromes, Alzheimer's and Huntington's diseases. It may also be used to reduce excess levels of prolactin in the treatment of galactorrhea and hypogonadism, and decrease the amount of free circulating somatostatin to prevent somatostatin's inhibitory effect on the release of growth hormone. The fusion of this protein to His-tag, HA-tag, IgG domains, etc. facilitates protein purification and fusion to IgG-1, IgG-3 and albumin increases the half life time in vivo.

Sequence 478 BP; 115 A; 121 C; 135 G; 107 T; 0 other;

Alignment Scores:
Pred. No.: 0.000459 Length: 478
Score: 112.00 Matches: 24
Percent Similarity: 53.57% Conservative: 6
Best Local Similarity: 42.86% Mismatches: 24
Query Match: 34.25% Indels: 2
DB: 20 Gaps: 2

09-833799-13B (1-57) x AAX80909 (1-478)
QY 1 AlaGlnGluProValLysGlyProValSerThrLysProGlySerCysProIleLeu 20
Db 76 GCTGTGGAAGGAGTTAAAGAGGGTATAGAG--AAAGCAGGGGTTTCCAGCTGACAAC 132
QY 21 IleArgCysAlaMetLeuAsnProProAsnArgCysLeuLysAspThrAspCysProGly 40
Db 133 GTACGCTGCTTCAAGTCCGATCCTCCC---CAGTGTACACAGACAGGACTGTCTGGGG 189
QY 41 IleLysLysCysCysGluGlySerCysGlyMetAlaCysPheValPro 56
Db 190 GAAAGGAAGTGTGTGTACCTGCACTGTGGCTTCAAGTGTGTGATTCCT 237

RESULT 67
AAX99904
ID AAA99904 standard; cDNA; 762 BP.

Qy	1	AlaGlnGluProValLysGlyProValSerThrLysProGlySerCysProIleLeu	20
		::	
Db	62	GCTGTGGAGGAGTTAAAGAGGGTATAGAG--AAAGCAGGGGTTTGCCAGCTGACAAC	118
Qy	21	IleArgCysAlaMetLeuAsnProAsnArgCysLeuLysAspThrAspCysProGly	40
		::: ::	
Db	119	GTACGCCTGCTCAAGTCCGATCCCTCCC---CAGTGTCCACACAGACCAGGACTGTCTGGGG	175
Qy	41	IleLysLysCysCysGluGlySerCysGlyMetalCysPheValpro	56

The method is based on changes in gene expression in BPH tissue isolated from patients exhibiting different clinical states of prostate hyperplasia as compared to normal prostate tissue. (I) comprises detecting the expression levels of one or more genes in prostate cells from the subject that are differentially regulated compared to normal prostate cells. (II) comprises preparing a first gene expression profile of BPH cells or BPH-like cell population, exposing the cells to the agent, preparing a second gene expression profile of the agent exposed cells, and comparing the first and second gene expression profiles. (I) is useful for diagnosing the onset or progression of BPH. (II) is useful for identifying an agent that modulates the onset or progression of BPH. The methods are useful to present information identifying the expression level in a tissue or cells, by comparing the expression level of genes given in the specification in the tissue or cells to the level of expression of gene in the database, and displaying the expression levels of at least one gene in the tissue or cell sample compared to the expression level in BPH. Agents using (II) are useful for treating BPH or prostate cancer. ABK64106-ABK64860 represent human benign prostatic hyperplasia gene sequences of the invention.

Sequence 411 BP; 98 A; 97 C; 94 G; 94 T; 28 other;

Alignment Scores:		
Pred. No.:	0.00144	411
Score:	107.00	21
Percent Similarity:	48.00%	3
Best Local Similarity:	42.00%	25
Query Match:	32.72%	1
DB:	24	0
	Length:	
	Matches:	
	Conservative:	
	Mismatches:	
	Indels:	
	Gaps:	

09-833799-13B (1-57) x ABK64114 (1-411)

[illegible]

RESULT 71

ABL66677/C

ID ABL66677 standard: DNA: 411 BP.

AC ABL66677:

DT 15-MAY-2002 (first entry)

Lung cancer related gene sequence SEO ID NO:5014.

Human; cancer; colon; breast; ovary; oesophagus; kidney; thyroid;
 KW stomach; lung; prostate; pancreas; carcinoma; antitumour; cancerous;
 KW cytostatic; gene therapy; antineoplastic; Wilm's tumour; adenocarcinoma;
 KW gene; ds.

OS Homo sapiens.

PN WO200194629-A2

XX
PD 13-DEC-2001

XX
PF 30-MAY-2001: 2001WQ-IIS10838

PR 05-JUN-2000: 2000US-209473P.

PR 05-JUN-2000; 2000US-209531P.

PR 18-SEP-2000; 2000US-233133P.

PR 18-SEP-2000; 2000US-233617P.

PR 20-SEP-2000; 2000US-234009P.

PR 20-SEP-2000; 2000US-234034P.

PR	20-SEP-2000;	2000US-2340502P.
PR	22-SEP-2000;	2000US-234509P.
PR	22-SEP-2000;	2000US-234567P.
PR	25-SEP-2000;	2000US-234923P.
PR	25-SEP-2000;	2000US-234924P.
PR	25-SEP-2000;	2000US-235077P.
PR	25-SEP-2000;	2000US-235082P.
PR	25-SEP-2000;	2000US-235134P.
PR	25-SEP-2000;	2000US-235280P.
PR	26-SEP-2000;	2000US-235637P.
PR	26-SEP-2000;	2000US-235638P.
PR	27-SEP-2000;	2000US-235711P.
PR	27-SEP-2000;	2000US-235720P.
PR	27-SEP-2000;	2000US-235840P.
PR	27-SEP-2000;	2000US-235863P.
PR	28-SEP-2000;	2000US-236028P.
PR	28-SEP-2000;	2000US-236032P.
PR	28-SEP-2000;	2000US-236033P.
PR	28-SEP-2000;	2000US-236034P.
PR	28-SEP-2000;	2000US-236109P.
PR	28-SEP-2000;	2000US-236111P.
PR	29-SEP-2000;	2000US-236842P.
PR	29-SEP-2000;	2000US-236891P.
PR	02-OCT-2000;	2000US-237172P.
PR	02-OCT-2000;	2000US-237173P.
PR	02-OCT-2000;	2000US-237278P.
PR	02-OCT-2000;	2000US-237294P.
PR	02-OCT-2000;	2000US-237295P.
PR	02-OCT-2000;	2000US-237316P.
PR	03-OCT-2000;	2000US-237425P.
PR	03-OCT-2000;	2000US-237598P.
PR	03-OCT-2000;	2000US-237604P.
PR	03-OCT-2000;	2000US-237606P.
PR	03-OCT-2000;	2000US-237608P.
PR	01-NOV-2000;	2000US-244867P.
PR	01-NOV-2000;	2000US-245084P.
XX		
PA	(AVAL-) AVALON PHARM.	
XX		
PI	Young PE, Augustus M, Cart	
PI	Soppet DR, Weaver Z;	
XX		
DR	WPI: 2002-188264/24.	

Screening for anti-neoplastic agent involves exposing cells to a chemical agent to be tested for anti-neoplastic activity, and determining a change in expression of a gene of a signature gene set.

Claim 1: SEO ID 5014: 44pp: English:

The present invention describes a method (M1) for screening for an anti-neoplastic agent. The method involves exposing cells to a chemical agent to be tested for anti-neoplastic activity, determining a change in expression of at least one gene (I) of a signature gene set, where (I) comprises a sequence (S) selected from 8447 sequences (given in ABL61664 to ABL70110), or is at least 95% identical to (S), where a change in expression is indicative of anti-neoplastic activity. (I) has cytostatic activity and can be used in gene therapy. M1 can be used for screening an anti-neoplastic agent, and can be used for producing a product which is the data collected with respect to the anti-neoplastic agent as a result of M1, and the data is sufficient to convey the chemical structure and/or properties of the agent. M1 can be used in the treatment of cancer such as colon, breast, stomach, lung, thyroid, oesophageal, ovarian, kidney, prostate or pancreatic cancer, adenocarcinoma, carcinoma, clear cell cancer, infiltrating ductal cancer, infiltrating lobular cancer, squamous cell carcinoma, neuroendocrine carcinoma, papillary carcinoma and Wilm's tumour.

Sequence 411 BP: 98 A: 97 C: 94 G: 94 T: 28 other:

Alignment Scores:

Alignment Scores:			
Pred. No.:	0.00144	Length:	411
Score:	107 00	Matches:	21

Percent Similarity: 48.00% Conservative: 3
Best Local Similarity: 42.00% Mismatches: 25
Query Match: 32.72% Indels: 1
DB: 24 Gaps: 0

09-833799-13B (1-57) x ABL66677 (1-411)

Qy 8 ProValSerThrLysProGlySerCysPro-IleIleLeuIleArgCysAlaMetLeuAs 27
Db 341 CCAACAAGGNGNAACCTTGGNAAGTCCNAGTGANTTNTGGCAATGTTNGATGCTNAA 282

Qy 27 nProProAsnArgCysLeuLysAspThrAspCysProGlyIleLysLysCysCysGluG1 47
Db 281 CCCCCCAATTCTGTGAGATGGATGGCCAGTGCAAGCGTGACTTGAAGTGTTCATGGG 222

Qy 47 ySerCysGlyMetAlaCysPheValPro 56
Db 221 CATGTGTGGAAATCCTGCGTTTCCCT 194

RESULT 72
ABL67461/c
ID ABL67461 standard; DNA; 411 BP.
XX
AC ABL67461;
XX
DT 15-MAY-2002 (first entry)
XX
DE Thyroid cancer related gene sequence SEQ ID NO:5798.
XX
KW Human; cancer; colon; breast; ovary; oesophagus; kidney; thyroid;
KW stomach; lung; prostate; pancreas; carcinoma; antitumour; cancerous;
KW cytostatic; gene therapy; antineoplastic; Wilm's tumour; adenocarcinoma;
KW gene; ds.
XX
OS Homo sapiens.
XX
PN WQ200194629-A2.
XX
PD 13-DEC-2001.
XX
PF 30-MAY-2001; 2001WO-US10838.
XX
PR 05-JUN-2000; 2000US-209473P.
PR 05-JUN-2000; 2000US-209531P.
PR 18-SEP-2000; 2000US-233133P.
PR 18-SEP-2000; 2000US-233617P.
PR 20-SEP-2000; 2000US-234009P.
PR 20-SEP-2000; 2000US-234034P.
PR 20-SEP-2000; 2000US-234052P.
PR 22-SEP-2000; 2000US-234509P.
PR 22-SEP-2000; 2000US-234567P.
PR 25-SEP-2000; 2000US-234923P.
PR 25-SEP-2000; 2000US-234924P.
PR 25-SEP-2000; 2000US-235077P.
PR 25-SEP-2000; 2000US-235082P.
PR 25-SEP-2000; 2000US-235134P.
PR 25-SEP-2000; 2000US-235280P.
PR 26-SEP-2000; 2000US-235637P.
PR 26-SEP-2000; 2000US-235638P.
PR 27-SEP-2000; 2000US-235711P.
PR 27-SEP-2000; 2000US-235720P.
PR 27-SEP-2000; 2000US-235840P.
PR 27-SEP-2000; 2000US-235863P.
PR 28-SEP-2000; 2000US-236028P.
PR 28-SEP-2000; 2000US-236032P.
PR 28-SEP-2000; 2000US-236033P.
PR 28-SEP-2000; 2000US-236034P.
PR 28-SEP-2000; 2000US-236109P.
PR 28-SEP-2000; 2000US-236111P.
PR 29-SEP-2000; 2000US-236842P.
PR 29-SEP-2000; 2000US-236891P.
PR 02-OCT-2000; 2000US-237172P.
PR 02-OCT-2000; 2000US-237173P.

PR 02-OCT-2000; 2000US-237278P.
PR 02-OCT-2000; 2000US-237294P.
PR 02-OCT-2000; 2000US-237295P.
PR 02-OCT-2000; 2000US-237316P.
PR 03-OCT-2000; 2000US-237425P.
PR 03-OCT-2000; 2000US-237598P.
PR 03-OCT-2000; 2000US-237604P.
PR 03-OCT-2000; 2000US-237606P.
PR 03-OCT-2000; 2000US-237608P.
PR 01-NOV-2000; 2000US-244867P.
PR 01-NOV-2000; 2000US-245084P.
XX
XX (AVAL-) AVALON PHARM.
XX
PI Young PE, Augustus M, Carter KC, Ebner R, Endress G, Horrigan S;
PI Soppet DR, Weaver Z;
XX
DR WPI; 2002-188264/24.
XX
PT Screening for anti-neoplastic agent involves exposing cells to a
PT chemical agent to be tested for anti-neoplastic activity, and
PT determining a change in expression of a gene of a signature gene set -
XX
PS Claim 1; SEQ ID 5798; 44pp; English.
XX
CC The present invention describes a method (M1) for screening for an
CC anti-neoplastic agent. The method involves exposing cells to a chemical
CC agent to be tested for anti-neoplastic activity, determining a change in
CC expression of at least one gene (I) of a signature gene set, where (I)
CC comprises a sequence (S) selected from 8447 sequences (given in ABL61664
CC to ABL70110), or is at least 95% identical to (S), where a change in
CC expression is indicative of anti-neoplastic activity. (I) has cytostatic
CC activity and can be used in gene therapy. M1 can be used for screening
CC an anti-neoplastic agent, and can be used for producing a product which
CC is the data collected with respect to the anti-neoplastic agent as a
CC result of M1, and the data is sufficient to convey the chemical
CC structure and/or properties of the agent. M1 can be used in the
CC treatment of cancer such as colon, breast, stomach, lung, thyroid,
CC oesophageal, ovarian, kidney, prostate or pancreatic cancer,
CC adenocarcinoma, carcinoma, clear cell cancer, infiltrating ductal cancer,
CC infiltrating lobular cancer, squamous cell carcinoma, neuroendocrine
CC carcinoma, papillary carcinoma and Wilm's tumour.
XX
SQ Sequence 411 BP; 98 A; 97 C; 94 G; 94 T; 28 other;

Alignment Scores:
Pred. No.: 0.00144 Length: 411
Score: 107.00 Matches: 21
Percent Similarity: 48.00% Conservative: 3
Best Local Similarity: 42.00% Mismatches: 25
Query Match: 32.72% Indels: 1
DB: 24 Gaps: 0

09-833799-13B (1-57) x ABL67461 (1-411)

Qy 8 ProValSerThrLysProGlySerCysPro-IleIleLeuIleArgCysAlaMetLeuAs 27
Db 341 CCAACAAGGNGNAACCTTGGNAAGTCCNAGTGANTTNTGGCAATGTTNGATGCTNAA 282

Qy 27 nProProAsnArgCysLeuLysAspThrAspCysProGlyIleLysLysCysCysGluG1 47
Db 281 CCCCCCAATTCTGTGAGATGGATGGCCAGTGCAAGCGTGACTTGAAGTGTTCATGGG 222

Qy 47 ySerCysGlyMetAlaCysPheValPro 56
Db 221 CATGTGTGGAAATCCTGCGTTTCCCT 194

RESULT 73
ABL67811/c
ID ABL67811 standard; DNA; 411 BP.
XX
AC ABL67811;
XX

DT 15-MAY-2002 (first entry)
XX Ovary cancer related gene sequence SEQ ID NO:6148.
DE Human; cancer; colon; breast; ovary; oesophagus; kidney; thyroid;
XX stomach; lung; prostate; pancreas; carcinoma; antitumour; cancerous;
KW cytostatic; gene therapy; antineoplastic; Wilm's tumour; adenocarcinoma;
KW gene; ds.
XX Homo sapiens.
OS WO200194629-A2.
XX 13-DEC-2001.
PD 30-MAY-2001; 2001WO-US10838.
XX 05-JUN-2000; 2000US-209473P.
PR 05-JUN-2000; 2000US-209531P.
PR 18-SEP-2000; 2000US-233133P.
PR 18-SEP-2000; 2000US-233617P.
PR 20-SEP-2000; 2000US-234009P.
PR 20-SEP-2000; 2000US-234034P.
PR 20-SEP-2000; 2000US-234052P.
PR 22-SEP-2000; 2000US-234509P.
PR 22-SEP-2000; 2000US-234567P.
PR 25-SEP-2000; 2000US-234923P.
PR 25-SEP-2000; 2000US-234924P.
PR 25-SEP-2000; 2000US-235077P.
PR 25-SEP-2000; 2000US-235082P.
PR 25-SEP-2000; 2000US-235134P.
PR 25-SEP-2000; 2000US-235280P.
PR 26-SEP-2000; 2000US-235637P.
PR 26-SEP-2000; 2000US-235638P.
PR 27-SEP-2000; 2000US-235711P.
PR 27-SEP-2000; 2000US-235720P.
PR 27-SEP-2000; 2000US-235840P.
PR 27-SEP-2000; 2000US-235863P.
PR 28-SEP-2000; 2000US-236028P.
PR 28-SEP-2000; 2000US-236032P.
PR 28-SEP-2000; 2000US-236033P.
PR 28-SEP-2000; 2000US-236034P.
PR 28-SEP-2000; 2000US-236109P.
PR 28-SEP-2000; 2000US-236111P.
PR 29-SEP-2000; 2000US-236842P.
PR 29-SEP-2000; 2000US-236891P.
PR 02-OCT-2000; 2000US-237172P.
PR 02-OCT-2000; 2000US-237173P.
PR 02-OCT-2000; 2000US-237278P.
PR 02-OCT-2000; 2000US-237294P.
PR 02-OCT-2000; 2000US-237295P.
PR 02-OCT-2000; 2000US-237316P.
PR 03-OCT-2000; 2000US-237425P.
PR 03-OCT-2000; 2000US-237598P.
PR 03-OCT-2000; 2000US-237604P.
PR 03-OCT-2000; 2000US-237606P.
PR 03-OCT-2000; 2000US-237608P.
PR 01-NOV-2000; 2000US-244867P.
PR 01-NOV-2000; 2000US-245084P.
XX (AVAL-) AVALON PHARM.
PA Young PE, Augustus M, Carter KC, Ebner R, Endress G, Horrigan S;
PI Soppet DR, Weaver Z;
XX WPI; 2002-188264/24.
XX Screening for anti-neoplastic agent involves exposing cells to a
PT chemical agent to be tested for anti-neoplastic activity, and
PT determining a change in expression of a gene of a signature gene set
XX Claim 1; SEQ ID 6148; 44pp; English.

CC The present invention describes a method (M1) for screening for an
CC anti-neoplastic agent. The method involves exposing cells to a chemical
CC agent to be tested for anti-neoplastic activity, determining a change in
CC expression of at least one gene (I) of a signature gene set, where (I)
CC comprises a sequence (S) selected from 8447 sequences (given in ABL61664
CC to ABL70110), or is at least 95% identical to (S), where a change in
CC expression is indicative of anti-neoplastic activity. (I) has cytostatic
CC activity and can be used in gene therapy. M1 can be used for screening
CC an anti-neoplastic agent, and can be used for producing a product which
CC is the data collected with respect to the anti-neoplastic agent as a
CC result of M1, and the data is sufficient to convey the chemical
CC structure and/or properties of the agent. M1 can be used in the
CC treatment of cancer such as colon, breast, stomach, lung, thyroid,
CC oesophageal, ovarian, kidney, prostate or pancreatic cancer,
CC adenocarcinoma, carcinoma, clear cell cancer, infiltrating ductal cancer,
CC infiltrating lobular cancer, squamous cell carcinoma, neuroendocrine
CC carcinoma, papillary carcinoma and Wilm's tumour.
XX
SQ Sequence 411 BP; 98 A; 97 C; 94 G; 94 T; 28 other;

Alignment Scores:
Pred. No.: 0.00144 Length: 411
Score: 107.00 Matches: 21
Percent Similarity: 48.00% Conservative: 3
Best Local Similarity: 42.00% Mismatches: 25
Query Match: 32.72% Indels: 1
DB: 24 Gaps: 0

09-833799-13B (1-57) x ABL67811 (1-411)

QY 8 ProValSerThrLysProGlySerCysPro-IleIleLeuIleArgCysAlaMetLeuAs 27
Db 341 CCAACAAGGNGNAACCCCTGGNAAGTNCAGTANTTNGCCCAATGTTNGATGCTNAA 282
QY 27 nProProAsnArgCysLeuLysAspThrAspCysProGlyIleLysLysCysCysGluG1 47
Db 281 CCCCCCAATTCTGTGAGATGGATGGCCAGTGCACGCGTGAAGTGTTCATGGG 222
QY 47 ySerCysGlyMetAlaCysPheValPro 56
Db 221 CATGTGTGGAAATCCTGCGTTTCCCCT 194

RESULT 74
AAA70066
ID AAA70066 standard; cDNA; 292 BP.
XX
AC AAA70066;
XX
DT 07-NOV-2000 (first entry)
XX Human ovarian carcinoma antigen polynucleotide SEQ ID NO:377.
DE Human; ovarian carcinoma; ovarian cancer; therapy; diagnosis;
KW tumour antigen; identification; cytostatic; gene therapy; vaccine; ss.
XX Homo sapiens.
OS WO200036107-A2.
XX
XX 22-JUN-2000.
PD
XX
XX 17-DEC-1999; 99WO-US30270.
XX
PR 17-DEC-1998; 98US-0215681.
PR 17-DEC-1998; 98US-0216003.
PR 23-JUN-1999; 99US-0338933.
PR 24-SEP-1999; 99US-0404879.
XX
PA (CORI-) CORIXA CORP.
XX Mitcham JL, King GE, Algate PA, Frudakis TN;
PI
XX WPI; 2000-431589/37.
DR

XX PT Immunogenic portion of an ovarian carcinoma protein and the nucleic
PT acid encoding it, useful for the diagnosis, prevention and treatment of
PT cancer, preferably ovarian cancer -
XX
PS Claim 18; Page 192; 299pp; English.
XX
CC The present invention describes an isolated polypeptide comprising an
CC immunogenic portion of an ovarian carcinoma protein (or its variants).
CC Ovarian carcinoma proteins, and polynucleotides encoding them, have
CC cytotstatic activity and can be used in gene therapy and vaccines.
CC Ovarian carcinoma polypeptides, nucleic acids, antibodies and vaccines
CC are useful for the prevention, diagnosis and treatment of cancer,
CC preferably ovarian cancer. AAA69691 to AAA70077 and AAB12552 to AAB12557
CC represent human ovarian carcinoma polynucleotides and proteins used in
CC the exemplification of the present invention.
XX
SQ Sequence 292 BP; 63 A; 71 C; 75 G; 83 T; 0 other;
Alignment Scores:
Pred. No.: 0.00161 Length: 292
Score: 105.00 Matches: 18
Percent Similarity: 57.14% Conservative: 2
Best Local Similarity: 51.43% Mismatches: 15
Query Match: 32.11% Indels: 0
DB: 21 Gaps: 0
09-833799-13B (1-57) x AAA70066 (1-292)
QY 22 ArgCysAlaMetLeuAsnProProAsnArgCysLeuLysAspThrAspCysProGlyIle 41
Db 1 CAATGTTTGATGCTTAACCCCCCAATTCTGTGAGATGGCCAGTGCAAGCGTGAC 60
QY 42 LysLysCysCysGluGlySerCysGlyMetAlaCysPheValPro 56
Db 61 TTGAAGTGTTCATGGGCATGTGTGGAAATCCTGCGTTTCCCT 105
RESULT 75
ABN72960
ID ABN72960 standard; DNA; 292 BP.
XX
AC ABN72960;
XX
DT 02-JUL-2002 (first entry)
XX
DE Ovarian carcinoma antigen polynucleotide #65.
XX
KW Human; immunostimulant; cytostatic; cancer; ovarian carcinoma; ds.
XX
OS Homo sapiens.
XX
PN WO200206317-A2.
XX
PD 24-JAN-2002.
XX
PF 17-JUL-2001; 2001WO-US22635.
XX
PR 17-JUL-2000; 2000US-0617747.
PR 10-AUG-2000; 2000US-0636801.
PR 20-SEP-2000; 2000US-0667857.
PR 04-APR-2001; 2001US-0827271.
PR 18-JUN-2001; 2001US-0884441.
XX
PA (CORI-) CORIXA CORP.
XX
PI Mitcham JL, King GE, Algate PA, Fling SP, Retter MW, Fanger GR;
PI Reed SG, Vedvick TS, Carter D, Hill P, Albone E;
XX
DR WPI; 2002-164781/21.
XX
PT Polypeptides comprising an immunogenic portion of an ovarian carcinoma
PT protein or its variants, useful for stimulating an immune response in a
PT patient and treating ovarian cancer -

XX
PS Example 2; Page 310; 408pp; English.
XX
CC This invention relates to polypeptides comprising an immunogenic
CC portion of an ovarian carcinoma protein which acts as an
CC immunostimulant and is cytostatic. The polypeptides, polynucleotides,
CC antibodies, fusion proteins, T cell populations and antigen presenting
CC cells that express the polypeptides are useful for stimulating an
CC immune response in a patient and treating ovarian cancer. This
CC sequence represents DNA related to the invention.
XX
SQ Sequence 292 BP; 63 A; 71 C; 75 G; 83 T; 0 other;
Alignment Scores:
Pred. No.: 0.00161 Length: 292
Score: 105.00 Matches: 18
Percent Similarity: 57.14% Conservative: 2
Best Local Similarity: 51.43% Mismatches: 15
Query Match: 32.11% Indels: 0
DB: 24 Gaps: 0
09-833799-13B (1-57) x ABN72960 (1-292)
QY 22 ArgCysAlaMetLeuAsnProProAsnArgCysLeuLysAspThrAspCysProGlyIle 41
Db 1 CAATGTTTGATGCTTAACCCCCCAATTCTGTGAGATGGCCAGTGCAAGCGTGAC 60
QY 42 LysLysCysCysGluGlySerCysGlyMetAlaCysPheValPro 56
Db 61 TTGAAGTGTTCATGGGCATGTGTGGAAATCCTGCGTTTCCCT 105
Search completed: February 15, 2003, 20:26:35
Job time : 232 secs

GenCore version 5.1.4 p5 4578
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: February 15, 2003, 20:11:56 ; Search time 1458 Seconds
(without alignments)
633.157 Million cell updates/sec

Title: 09-833799-13B
Perfect score: 327
Sequence: 1 aqepvkgpvstkpqscpiil.....cpgikkcccgscgmactfvpq 57

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 16154066 seqs, 8097743376 residues

Total number of hits satisfying chosen parameters: 32308132

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 100 summaries

Command line parameters:
-MODEL=frame+p2n.model -DEV=xlp
-Q=/cgn2_1/USPTO_spool/US09833799/runat_12022003_120042_23664/app_query.fasta_1.199
-DB=EST -QFMT=fastcap -SUFFIX=1st -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=-1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=100
-DOCALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=75 -MODE=LOCAL
-OUTFMT=pto -NORM=ext -HEAPSIZ=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US09833799@cgn 1 1 2874 @runat_12022003_120042_23664 -NCPU=6 -ICPU=3
-NO_XLPXY -NO_MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -LONGLOG -DEV TIMEOUT=120
-WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : EST:
1: em_estba:*
2: em_esthum:*
3: em_estin:*
4: em_estmu:*
5: em_estov:*
6: em_estpl:*
7: em_estro:*
8: em_htc:*
9: gb_est1:*
10: gb_est2:*
11: gb_htc:*
12: gb_est3:*
13: gb_est4:*
14: gb_est5:*
15: em_estfun:*
16: em_estom:*
17: gb_gss:*
18: em_gss_hum:*
19: em_gss_inv:*
20: em_gss_pln:*
21: em_gss_vrt:*
22: em_gss_fun:*
23: em_gss_mam:*
24: em_gss_mus:*
25: em_gss_other:*
26: em_gss_pro:*
27: em_gss_rod:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	DB	ID	Description
1	327	100.0	219	10	BE182652	BE182652 RC3-HT064
2	327	100.0	289	10	BE181373	BE181373 CM0-HT063
3	327	100.0	312	10	BE181959	BE181959 CM1-HT064
4	327	100.0	340	13	BG987761	BG987761 CM3-HT119
5	327	100.0	342	10	BE182993	BE182993 CM4-HT065
6	327	100.0	344	10	BE181546	BE181546 CM0-HT063
7	327	100.0	359	10	BE182864	BE182864 CM4-HT065
8	327	100.0	364	12	BE715602	BE715602 CM4-HT074
9	327	100.0	370	12	BE715605	BE715605 CM4-HT074
10	327	100.0	370	12	BE715626	BE715626 CM4-HT074
11	327	100.0	370	12	BE772160	BE772160 CM4-HT010
12	327	100.0	371	12	BF838555	BF838555 IL5-HT099
13	327	100.0	371	12	BE772166	BE772166 CM4-HT010
14	327	100.0	375	9	AI285400	AI285400 qt69f06.x
15	327	100.0	379	12	BE715603	BE715603 CM4-HT074
16	327	100.0	380	12	BE772155	BE772155 CM4-HT010
17	327	100.0	380	12	BE772171	BE772171 CM4-HT010
18	327	100.0	386	10	BE182879	BE182879 CM4-HT065
19	327	100.0	386	12	BE772169	BE772169 CM4-HT010
20	327	100.0	388	9	AA583997	AA583997 nn65a01.s
21	327	100.0	388	12	BE772159	BE772159 CM4-HT010
22	327	100.0	389	12	BE772168	BE772168 CM4-HT010
23	327	100.0	390	9	AI393896	AI393896 tg05e09.x
24	327	100.0	390	9	AI989737	AI989737 wu21e08.x
25	327	100.0	390	12	BE715611	BE715611 CM4-HT074
26	327	100.0	395	12	BE715615	BE715615 CM4-HT074
27	327	100.0	397	12	BE772137	BE772137 CM4-HT010
28	327	100.0	399	9	AA586974	AA586974 nn69d08.s
29	327	100.0	400	12	BE715592	BE715592 CM4-HT074
30	327	100.0	403	12	BE715609	BE715609 CM4-HT074
31	327	100.0	405	9	AI874186	AI874186 wt50a02.x
32	327	100.0	406	10	AW137392	AW137392 UI-H-B11-
33	327	100.0	406	12	BE715604	BE715604 CM4-HT074
34	327	100.0	418	9	AI392753	AI392753 tg23g03.x
35	327	100.0	419	9	AA586943	AA586943 nn69a01.s
36	327	100.0	422	10	AW082097	AW082097 xb60d08.x
37	327	100.0	428	12	BE772154	BE772154 CM4-HT010
38	327	100.0	429	9	AI283910	AI283910 qt66a10.x
39	327	100.0	437	12	BF836224	BF836224 QV3-HT101
40	327	100.0	438	10	AW001880	AW001880 wt81d06.x
41	327	100.0	438	12	BE715601	BE715601 CM4-HT074
42	327	100.0	441	12	BF836743	BF836743 CM2-HT096
43	327	100.0	448	9	AI582329	AI582329 tq66g09.x
44	327	100.0	453	9	AI459240	AI459240 tk11c09.x
45	327	100.0	465	12	BF824934	BF824934 IL0-HN003
46	327	100.0	466	12	BF824930	BF824930 IL0-HN003
47	327	100.0	466	12	BF824935	BF824935 IL0-HN003
48	327	100.0	467	12	BF824920	BF824920 IL0-HN003
49	327	100.0	467	12	BF825207	BF825207 IL0-HN003
50	327	100.0	471	12	BF837734	BF837734 QV3-HT101
51	327	100.0	477	12	BF824932	BF824932 IL0-HN003
52	327	100.0	480	9	AA582866	AA582866 nn72a09.s
53	327	100.0	487	12	BF825204	BF825204 IL0-HN003
54	327	100.0	490	12	BF002099	BF002099 7g99b04.x
55	327	100.0	498	12	BF837735	BF837735 QV3-HT101
56	327	100.0	522	14	BQ941085	BQ941085 AGENCOURT
57	327	100.0	565	12	BG490472	BG490472 602519610
58	324	99.1	361	12	BE772161	BE772161 CM4-HT010
59	324	99.1	421	13	BI061069	BI061069 IL3-UT011
60	323	98.8	263	12	BE711100	BE711100 RC3-HT064
61	323	98.8	420	9	AI924155	AI924155 wt64c08.x
62	323	98.8	447	12	BF835679	BF835679 QV3-HT101
63	323	98.8	453	10	AW845140	AW845140 IL0-CT000
64	323	98.8	455	10	AW845142	AW845142 IL0-CT000

```
c 65 322 98.5 444 9 AI858070
c 66 321 98.2 342 12 BE715619
c 67 321 98.2 374 12 BF836201
c 68 321 98.2 501 12 BF825196
c 69 320 97.9 316 12 BE711083
c 70 320 97.9 344 10 AW844969
c 71 320 97.9 428 14 BM768573
c 72 319 97.6 394 13 BI021805
c 73 318 97.2 386 9 AA586983
c 74 318 97.2 419 12 BF736734
c 75 318 97.2 423 9 AI830872
c 76 318 97.2 527 9 AI807596
c 77 316 96.6 431 12 BF837730
c 78 315 96.3 298 12 BF837690
c 79 315 96.3 313 12 BF736294
c 80 315 96.3 314 12 BF837677
c 81 315 96.3 345 12 BF094327
c 82 314 96.0 275 12 BE711090
c 83 314 96.0 391 12 BE772162
c 84 314 96.0 393 10 BE183902
c 85 314 96.0 393 12 BE772164
c 86 314 96.0 448 10 AW082007
c 87 313 95.7 467 12 BF836725
c 88 311 95.1 289 12 BF834601
c 89 309 94.5 275 10 BE182650
c 90 308 94.2 344 10 AW860590
c 91 307 93.9 445 9 AI830232
c 92 305 93.3 470 12 BF913869
c 93 302 92.4 400 9 AI831796
c 94 300 91.7 555 9 AA583567
c 95 298 91.1 411 9 AA586718
c 96 298 91.1 478 9 AA527557
c 97 291 89.0 190 10 BE181433
c 98 290.5 88.8 468 12 BF090376
c 99 286 87.5 299 10 BE184317
100 284 86.9 470 12 BF913875
```

ALIGNMENTS

```
RESULT 1
BE182652 219 bp mRNA linear EST 22-JUN-2000
LOCUS RC3-HT0649-090500-011-h03 HT0649 Homo sapiens cDNA, mRNA sequence.
DEFINITION BE182652
ACCESSION BE182652.1 GI:8661828
VERSION EST.
KEYWORDS human.
SOURCE Homo sapiens
ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
REFERENCE Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
AUTHORS 1 (bases 1 to 219)
Dias Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Briones,M.R.,
Nagai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.F.,
Goldman,G.H., Carvalho,A.F., Matsukuma,A., Baia,G.S., Simpson,D.H.,
Brunstein,A., deOliveira,P.S., Bucher,P., Jongeneel,C.V., O'Hare
,M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and
Simpson,A.J.
```

```
TITLE Shotgun sequencing of the human transcriptome with ORF expressed
sequence tags
JOURNAL Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
MEDLINE 20202663
COMMENT Contact: Simpson A.J.G.
Laboratory of Cancer Genetics
Ludwig Institute for Cancer Research
Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
Brazil
Tel: +55-11-2704922
Fax: +55-11-2707001
Email: asimpson@ludwig.org.br
This sequence was derived from the FAPESP/LICR Human Cancer Genome
Project. This entry can be seen in the following URL
(http://www.ludwig.org.br/scripts/gethtml2.pl?t1=&t2=RC3-HT0649-090
```

```
500-011-h03&t3=2000-05-09&t4=1)
Seq primer: puc 18 forward
High quality sequence start: 22
High quality sequence stop: 219.
Location/Qualifiers
```

FEATURES

source

```
1..219
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_lib="HT0649"
/dev_stage="Adult"
/note="Organ: head_neck; Vector: puc18; Site_1: SmaI;
Site_2: SmaI; A mini-library was made by cloning products
derived from ORESTES PCR (U.S. Letters Patent application
No. 196,716 - Ludwig Institute for Cancer Research)
profiles into the pUC 18 vector. Reverse transcription of
tissue mRNA and cDNA amplification were performed under
low stringency conditions."
```

```
BASE COUNT 51 a 60 c 59 g 49 t
ORIGIN
```

```
Alignment Scores: 1.44e-29 Length: 219
```

```
Pred. No.: 327.00 Matches: 57
```

```
Score: 100.00% Conservative: 0
```

```
Percent Similarity: 100.00% Mismatches: 0
```

```
Best Local Similarity: 100.00% Indels: 0
```

```
Query Match: 100.00% Gaps: 0
```

```
DB: 10
```

```
09-833799-13B (1-57) x BE182652 (1-219)
```

```
QY 1 AlaGlnGluProValLysGlyProValSerThrLysProGlySerCysProIleLeu 20
|||||
Db 24 GCGCAAGAGCCAGTCAAAGGTCCAGTCTCCACTAAGCTGGCTCCTGCCCATATCTTG 83
|||||
```

```
QY 21 IleArgCysAlaMetLeuAsnProProAsnArgCysLeuLysAspThrAspCysProGly 40
|||||
Db 84 ATCCGGTGGCCCATGTTGAATCCCCCTAACCGCTGCTGAAAGATACTGACTGCCAGGA 143
|||||
```

```
QY 41 IleLysLysCysCysGluGlySerCysGlyMetAlaCysPheValProGln 57
|||||
```

```
Db 144 ATCAAGAAGTGTGTGAAGGCTCTTGGGGATGGCCCTGTTTCGTTCCCCAG 194
|||||
```

RESULT 2

```
BE181373/c
```

```
LOCUS
```

```
DEFINITION CM0-HT0633-240300-304-g12 HT0633 Homo sapiens cDNA, mRNA sequence.
```

```
ACCESSION BE181373
```

```
VERSION BE181373.1 GI:8660549
```

```
KEYWORDS EST.
```

```
SOURCE human.
```

```
ORGANISM Homo sapiens
```

```
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
```

```
AUTHORS Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
```

```
1 (bases 1 to 289)
```

```
Dias Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Briones,M.R.,
```

```
Nagai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.F.,
```

```
Goldman,G.H., Carvalho,A.F., Matsukuma,A., Baia,G.S., Simpson,D.H.,
```

```
Brunstein,A., deOliveira,P.S., Bucher,P., Jongeneel,C.V., O'Hare
```

```
,M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and
```

```
Simpson,A.J.
```

```
Shotgun sequencing of the human transcriptome with ORF expressed
```

```
sequence tags
```

```
JOURNAL Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
```

```
MEDLINE 20202663
```

```
COMMENT Contact: Simpson A.J.G.
```

```
Laboratory of Cancer Genetics
```

```
Ludwig Institute for Cancer Research
```

```
Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
```

```
Brazil
```

```
Tel: +55-11-2704922
```

```
Fax: +55-11-2707001
```

```
Email: asimpson@ludwig.org.br
```

```
This sequence was derived from the FAPESP/LICR Human Cancer Genome
```

```
Project. This entry can be seen in the following URL
```

```
(http://www.ludwig.org.br/scripts/gethtml2.pl?t1=&t2=RC3-HT0649-090
```

Project. This entry can be seen in the following URL
(http://www.ludwig.org.br/scripts/gethtml2.pl?tl=&t2=CM0-HT0633-240
300-304-gl2&t3=2000-03-24&t4=1)
Seq primer: puc 18 forward
High quality sequence start: 17
High quality sequence stop: 289.

FEATURES
source

1. .289
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_lib="HT0633"
/dev_stage="Adult"
/note="Organ: head_neck; Vector: puc18; Site_1: SmaI;
Site_2: SmaI; A mini-library was made by cloning products
derived from ORESTES PCR (U.S. Letters Patent application
No. 196,716 - Ludwig Institute for Cancer Research)
profiles into the pUC 18 vector. Reverse transcription of
tissue mRNA and cDNA amplification were performed under
low stringency conditions."

BASE COUNT 62 a 74 c 93 g 60 t
ORIGIN

Alignment Scores:
Pred. No.: 1.99e-29 Length: 289
Score: 327.00 Matches: 57
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 10 Gaps: 0

09-833799-13B (1-57) x BE181373 (1-289)

QY 1 AlaGlnGluProValSerThrLysProGlySerCysProIleLeu 20
|||||
Db 274 GCGCAGGAGCCAGTCAAGGTCAGTCTCCACTAAGCCTGGCTCCTGCCCATATCTTG 215
|||||
QY 21 IleArgCysAlaMetLeuAsnProProAsnArgCysLeuLysAspThrAspCysProGly 40
|||||
Db 214 ATCCGGTGGCCCATGTTGAATCCCCCTAACCGCTGCTTGAAAGATACTGACTGCCCAGGA 155
|||||
QY 41 IleLysLysCysCysGluGlySerCysGlyMetAlaCysPheValProGln 57
|||||
Db 154 ATCAAGAAAGTGCTGTGAAGGCTCTTGCGGGATGGCCTGTTTCGTTCCCCAG 104
|||||

RESULT 3
BE181959
LOCUS BE181959 312 bp mRNA linear EST 22-JUN-2000
DEFINITION CM1-HT0641-280400-206-d02 HT0641 Homo sapiens cDNA, mRNA sequence.
ACCESSION BE181959
VERSION BE181959.1 GI:8661135
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens

REFERENCE
AUTHORS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 312)
Dias Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Briones,M.R.,
Nagai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.F.,
Goldman,G.H., Carvalho,A.F., Matsukuma,A., Baia,G.S., Simpson,D.H.,
Brunstein,A., deOliveira,P.S., Bucher,P., Jongeneel,C.V., O'Hare
M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and
Simpson,A.J.

TITLE Shotgun sequencing of the human transcriptome with ORF expressed
sequence tags
JOURNAL Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
MEDLINE 20202663
COMMENT Contact: Simpson A.J.G.
Laboratory of Cancer Genetics
Ludwig Institute for Cancer Research
Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
Brazil
Tel: +55-11-2704922
Fax: +55-11-2707001

Email: asimpson@ludwig.org.br
This sequence was derived from the FAPESP/LICR Human Cancer Genome
Project. This entry can be seen in the following URL
(http://www.ludwig.org.br/scripts/gethtml2.pl?tl=&t2=CM1-HT0641-280
400-206-d02&t3=2000-04-28&t4=1)
Seq primer: puc 18 forward
High quality sequence start: 15
High quality sequence stop: 312.

FEATURES
source

1. .312
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_lib="HT0641"
/dev_stage="Adult"
/note="Organ: head_neck; Vector: puc18; Site_1: SmaI;
Site_2: SmaI; A mini-library was made by cloning products
derived from ORESTES PCR (U.S. Letters Patent application
No. 196,716 - Ludwig Institute for Cancer Research)
profiles into the pUC 18 vector. Reverse transcription of
tissue mRNA and cDNA amplification were performed under
low stringency conditions."

BASE COUNT 70 a 93 c 81 g 68 t
ORIGIN

Alignment Scores:
Pred. No.: 2.18e-29 Length: 312
Score: 327.00 Matches: 57
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 10 Gaps: 0

09-833799-13B (1-57) x BE181959 (1-312)

QY 1 AlaGlnGluProValSerThrLysProGlySerCysProIleLeu 20
|||||
Db 71 GCGCAGAGCCGGTCAAGGTCAGTCTCCACTAAGCCTGGCTCCTGCCCATATCTTG 130
|||||
QY 21 IleArgCysAlaMetLeuAsnProProAsnArgCysLeuLysAspThrAspCysProGly 40
|||||
Db 131 ATCCGGTGGCCCATGTTGAATCCCCCTAACCGCTGCTTGAAAGATACTGACTGCCCAGGA 190
|||||
QY 41 IleLysLysCysCysGluGlySerCysGlyMetAlaCysPheValProGln 57
|||||
Db 191 ATCAAGAAAGTGCTGTGAAGGCTCTTGCGGGATGGCCTGTTTCGTTCCCCAG 241
|||||

RESULT 4
BG987761/c
LOCUS BG987761 340 bp mRNA linear EST 13-JUN-2001
DEFINITION CM3-HT1190-110101-624-h04 HT1190 Homo sapiens cDNA, mRNA sequence.
ACCESSION BG987761
VERSION BG987761.1 GI:14391831
KEYWORDS EST.
SOURCE human.

ORGANISM Homo sapiens
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 340)
Dias Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Briones,M.R.,
Nagai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.F.,
Goldman,G.H., Carvalho,A.F., Matsukuma,A., Baia,G.S., Simpson,D.H.,
Brunstein,A., deOliveira,P.S., Bucher,P., Jongeneel,C.V., O'Hare
M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and
Simpson,A.J.

TITLE Shotgun sequencing of the human transcriptome with ORF expressed
sequence tags
JOURNAL Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
MEDLINE 20202663
COMMENT Contact: Simpson A.J.G.
Laboratory of Cancer Genetics
Ludwig Institute for Cancer Research
Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
Brazil

Tel: +55-11-2704922
Fax: +55-11-2707001
Email: asimpson@ludwig.org.br
This sequence was derived from the FAPESP/LICR Human Cancer Genome Project. This entry can be seen in the following URL
(http://www.ludwig.org.br/scripts/gethtml2.pl?tl=CM3&t2=CM3-HT1190-110101-624-h04&t3=2001-01-11&t4=1)
Seq primer: puc 18 forward
High quality sequence start: 6
High quality sequence stop: 340.
Location/Qualifiers

FEATURES

source
1. .340
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_lib="HT1190"
/dev_stage="Adult"
/note="Organ: head neck; Vector: puc18; Site 1: SmaI;
Site 2: SmaI; A mini-library was made by cloning products
derived from ORESTES PCR (U.S. Letters Patent application
No. 196,716 - Ludwig Institute for Cancer Research)
profiles into the pUC 18 vector. Reverse transcription of
tissue mRNA and cDNA amplification were performed under
low stringency conditions."

BASE COUNT 79 a 82 c 97 g 82 t
ORIGIN

Alignment Scores: 2.41e-29 Length: 340
Pred. No.: 327.00 Matches: 57
Score: 100.00% Conservative: 0
Percent Similarity: 100.00% Mismatches: 0
Best Local Similarity: 100.00% Indels: 0
Query Match: 13 Gaps: 0
DB:

09-833799-13B (1-57) x BG987761 (1-340)

QY 1 AlaGlnGluProValLysGlyProValSerThrLysProGlySerCysProIleLeu 20
|||||
Db 238 GCGCAAGAGCCAGTCAAAGGTCCTCCACTAAGCCTGGCTCCTGCCCATTTATCTG 179
QY 21 IleArgCysAlaMetLeuAsnProProAsnArgCysLeuLysAspThrAspCysProGly 40
|||||
Db 178 ATCCGGTGGCCATGTTGAATCCCCCTAACCGCTGCTTGAAGATACCTGACTGCCACGGA 119
QY 41 IleLysLysCysCysGluGlySerCysGlyMetAlaCysPheValProGln 57
|||||
Db 118 ATCAAGAAGTGTGTGAAGGCTCTTGGCGGATGGCCTGTTTCGTTCCCCAG 68

RESULT 5
BE182993/c 342 bp mRNA linear EST 22-JUN-2000

LOCUS BE182993
DEFINITION CM4-HT0653-180400-146-all HT0653 Homo sapiens cDNA, mRNA sequence.
ACCESSION BE182993
VERSION BE182993.1 GI:8662169
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 342)
AUTHORS Dias Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Briones,M.R.,
Nagai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.F.,
Goldman,G.H., Carvalho,A.F., Matsukuma,A., Baia,G.S., Simpson,D.H.,
Brunstein,A., deOliveira,P.S., Bucher,P., Jongeneel,C.V., O'Hare
M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and
Simpson,A.J.

TITLE Shotgun sequencing of the human transcriptome with ORF expressed
sequence tags

JOURNAL Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
MEDLINE 20202663
COMMENT Contact: Simpson A.J.G.
Laboratory of Cancer Genetics
Ludwig Institute for Cancer Research

Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
Brazil
Tel: +55-11-2704922
Fax: +55-11-2707001
Email: asimpson@ludwig.org.br
This sequence was derived from the FAPESP/LICR Human Cancer Genome Project. This entry can be seen in the following URL
(http://www.ludwig.org.br/scripts/gethtml2.pl?tl=&t2=CM4-HT0653-180400-146-all&t3=2000-04-18&t4=1)
Seq primer: puc 18 forward
High quality sequence start: 21
High quality sequence stop: 342.
Location/Qualifiers

FEATURES

source
1. .342
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_lib="HT0653"
/dev_stage="Adult"
/note="Organ: head neck; Vector: puc18; Site 1: SmaI;
Site 2: SmaI; A mini-library was made by cloning products
derived from ORESTES PCR (U.S. Letters Patent application
No. 196,716 - Ludwig Institute for Cancer Research)
profiles into the pUC 18 vector. Reverse transcription of
tissue mRNA and cDNA amplification were performed under
low stringency conditions."

BASE COUNT 80 a 82 c 100 g 79 t 1 others
ORIGIN

Alignment Scores: 2.43e-29 Length: 342
Pred. No.: 327.00 Matches: 57
Score: 100.00% Conservative: 0
Percent Similarity: 100.00% Mismatches: 0
Best Local Similarity: 100.00% Indels: 0
Query Match: 10 Gaps: 0
DB:

09-833799-13B (1-57) x BE182993 (1-342)

QY 1 AlaGlnGluProValLysGlyProValSerThrLysProGlySerCysProIleLeu 20
|||||
Db 259 GCGCAAGAGCCAGTCAAAGGTCCTCCACTAAGCCTGGCTCCTGCCCATTTATCTG 200
QY 21 IleArgCysAlaMetLeuAsnProProAsnArgCysLeuLysAspThrAspCysProGly 40
|||||
Db 199 ATCCGGTGGCCATGTTGAATCCCCCTAACCGCTGCTTGAAGATACCTGACTGCCACGGA 140
QY 41 IleLysLysCysCysGluGlySerCysGlyMetAlaCysPheValProGln 57
|||||
Db 139 ATCAAGAAGTGTGTGAAGGCTCTTGGCGGATGGCCTGTTTCGTTCCCCAG 89

RESULT 6
BE181546/c 344 bp mRNA linear EST 22-JUN-2000

LOCUS BE181546
DEFINITION CM0-HT0634-270300-308-f02 HT0634 Homo sapiens cDNA, mRNA sequence.
ACCESSION BE181546
VERSION BE181546.1 GI:8660722
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 344)
AUTHORS Dias Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Briones,M.R.,
Nagai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.F.,
Goldman,G.H., Carvalho,A.F., Matsukuma,A., Baia,G.S., Simpson,D.H.,
Brunstein,A., deOliveira,P.S., Bucher,P., Jongeneel,C.V., O'Hare
M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and
Simpson,A.J.

TITLE Shotgun sequencing of the human transcriptome with ORF expressed
sequence tags

JOURNAL Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
MEDLINE 20202663
COMMENT Contact: Simpson A.J.G.

Laboratory of Cancer Genetics
Ludwig Institute for Cancer Research
Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
Brazil
Tel: +55-11-2704922
Fax: +55-11-2707001
Email: asimpson@ludwig.org.br
This sequence was derived from the FAPESP/LICR Human Cancer Genome
Project. This entry can be seen in the following URL
(http://www.ludwig.org.br/scripts/gethtml2.pl?tl=&t2=CM0-HT0634-270
300-308-f02&t3=2000-03-27&t4=1)
Seq primer: puc 18 forward
High quality sequence start: 6
High quality sequence stop: 344.
Location/Qualifiers

FEATURES
source

1. .344
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_lib="HT0634"
/dev_stage="Adult"
/note="Organ: head neck; Vector: puc18; Site: 1: SmaI;
Site 2: SmaI; A mini-library was made by cloning products
derived from ORESTES PCR (U.S. Letters Patent application
No. 196,716 - Ludwig Institute for Cancer Research)
profiles into the pUC 18 vector. Reverse transcription of
tissue mRNA and cDNA amplification were performed under
low stringency conditions."
79 a 80 c 113 g 72 t

BASE COUNT 79 a 80 c 113 g 72 t
ORIGIN
Alignment Scores:
Pred. No.: 2.45e-29 Length: 344
Score: 327.00 Matches: 57
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 10 Gaps: 0

09-833799-13B (1-57) x BE181546 (1-344)

Qy 1 AlaGlnGluProValLysGlyProValSerThrLysProGlySerCysProIleLeu 20
|||||
Db 294 GCGCAAGAGCCAGTCAAGAGTCCAGTCTCCACTAAGCTGGCTCTGCCCCATTATCTTG 235
Qy 21 IleArgCysAlaMetLeuAsnProProAsnArgCysLeuLysAspThrAspCysProGly 40
|||||
Db 234 ATCCGGTGGCCCATGTTGAATCCCCCTAACCGCTGCTTGAAAGATACTGACTGCCAGGA 175
Qy 41 IleLysLysCysCysGluGlySerCysGlyMetAlaCysPheValProGln 57
|||||
Db 174 ATCAAGAAGTGCTGTGAAGGCTCTTGCGGGATGGCTGTTTCGTTCCCCAG 124

RESULT 7
BE182864
LOCUS BE182864 359 bp mRNA linear EST 22-JUN-2000
DEFINITION CM4-HT0652-150400-143-a07 HT0652 Homo sapiens cDNA, mRNA sequence.
ACCESSION BE182864
VERSION BE182864.1 GI:8662040
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 359)
AUTHORS Dias Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Briones,M.R.,
Nagai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.F.,
Goldman,G.H., Carvalho,A.F., Matsukuma,A., Baia,G.S., Simpson,D.H.,
Brunstein,A., deOliveira,P.S., Bucher,P., Jongeneel,C.V., O'Hare
M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and
Simpson,A.J.
TITLE Shotgun sequencing of the human transcriptome with ORF expressed
sequence tags
JOURNAL Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)

MEDLINE
COMMENT

20202663
Contact: Simpson A.J.G.
Laboratory of Cancer Genetics
Ludwig Institute for Cancer Research
Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
Brazil
Tel: +55-11-2704922
Fax: +55-11-2707001
Email: asimpson@ludwig.org.br
This sequence was derived from the FAPESP/LICR Human Cancer Genome
Project. This entry can be seen in the following URL
(http://www.ludwig.org.br/scripts/gethtml2.pl?tl=&t2=CM4-HT0652-150
400-143-a07&t3=2000-04-15&t4=1)
Seq primer: puc 18 forward
High quality sequence start: 39
High quality sequence stop: 359.
Location/Qualifiers

FEATURES
source

1. .359
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_lib="HT0652"
/dev_stage="Adult"
/note="Organ: head neck; Vector: puc18; Site: 1: SmaI;
Site 2: SmaI; A mini-library was made by cloning products
derived from ORESTES PCR (U.S. Letters Patent application
No. 196,716 - Ludwig Institute for Cancer Research)
profiles into the pUC 18 vector. Reverse transcription of
tissue mRNA and cDNA amplification were performed under
low stringency conditions."
92 a 90 c 95 g 82 t

BASE COUNT 92 a 90 c 95 g 82 t
ORIGIN
Alignment Scores:
Pred. No.: 2.57e-29 Length: 359
Score: 327.00 Matches: 57
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 10 Gaps: 0

09-833799-13B (1-57) x BE182864 (1-359)

Qy 1 AlaGlnGluProValLysGlyProValSerThrLysProGlySerCysProIleLeu 20
|||||
Db 172 GCGCAAGAGCCAGTCAAGAGTCCAGTCTCCACTAAGCTGGCTCTGCCCCATTATCTTG 231
Qy 21 IleArgCysAlaMetLeuAsnProProAsnArgCysLeuLysAspThrAspCysProGly 40
|||||
Db 232 ATCCGGTGGCCCATGTTGAATCCCCCTAACCGCTGCTTGAAAGATACTGACTGCCAGGA 291
Qy 41 IleLysLysCysCysGluGlySerCysGlyMetAlaCysPheValProGln 57
|||||
Db 292 ATCAAGAAGTGCTGTGAAGGCTCTTGCGGGATGGCTGTTTCGTTCCCCAG 342

RESULT 8
BE182864/c
LOCUS BE182864 364 bp mRNA linear EST 12-SEP-2000
DEFINITION CM4-HT0744-160600-201-a06 HT0744 Homo sapiens cDNA, mRNA sequence.
ACCESSION BE182864
VERSION BE182864.1 GI:10103867
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 364)
AUTHORS Dias Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Briones,M.R.,
Nagai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.F.,
Goldman,G.H., Carvalho,A.F., Matsukuma,A., Baia,G.S., Simpson,D.H.,
Brunstein,A., deOliveira,P.S., Bucher,P., Jongeneel,C.V., O'Hare
M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and
Simpson,A.J.
TITLE Shotgun sequencing of the human transcriptome with ORF expressed

JOURNAL
MEDLINE
COMMENT

sequence tags
Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
20202663
Contact: Simpson A.J.G.
Laboratory of Cancer Genetics
Ludwig Institute for Cancer Research
Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
Brazil
Tel: +55-11-2704922
Fax: +55-11-2707001
Email: asimpson@ludwig.org.br
This sequence was derived from the FAPESP/LICR Human Cancer Genome
Project. This entry can be seen in the following URL
(http://www.ludwig.org.br/scripts/gethtml2.pl?t1=&t2=CM4-HT0744-160
600-201-a06&t3=2000-06-16&t4=1)
Seq primer: puc 18 forward
High quality sequence start: 20
High quality sequence stop: 364.

FEATURES
source

1. .364
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_lib="HT0744"
/dev_stage="Adult"
/note="Organ: head neck; Vector: puc18; Site_1: SmaI;
Site_2: SmaI; A mini-library was made by cloning products
derived from ORESTES PCR (U.S. Letters Patent application
No. 196,716 - Ludwig Institute for Cancer Research)
profiles into the pUC 18 vector. Reverse transcription of
tissue mRNA and cDNA amplification were performed under
low stringency conditions."
BASE COUNT 85 a 92 c 100 g 87 t
ORIGIN

Alignment Scores:
Pred. No.: 2.61e-29 Length: 364
Score: 327.00 Matches: 57
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 12 Gaps: 0

09-833799-13B (1-57) x BE715602 (1-364)

Qy 1 AlaGlnGluProValLysGlyProValSerThrLysProGlySerCysProIleLeu 20
Db 239 GCGCAAGAGCCAGTCAAAGGTCAGTCTCCACTAAGCGTGGCTCCTGCCCATATCTTG 180
Qy 21 IleArgCysAlaMetLeuAsnProProAsnArgCysLeuLysAspThrAspCysProGly 40
Db 179 ATCCGGTGGCCCATGTTGAATCCCCCTAACCGCTGCTTGAAGAGATACCTGACTGCCCAGGA 120
Qy 41 IleLysLysCysCysGluGlySerCysGlyMetAlaCysPheValProGln 57
Db 119 ATCAAGAAGTGCTGTGAAGGCTCTTGGGGATGGCCTGTTTCGTTCCCCAG 69

RESULT 9
BE715605/c

LOCUS BE715605 370 bp mRNA linear EST 12-SEP-2000
DEFINITION CM4-HT0744-160600-201-a10 HT0744 Homo sapiens cDNA, mRNA sequence.
ACCESSION BE715605
VERSION BE715605.1 GI:10103870
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 370)
AUTHORS Dias Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Briones,M.R.,
Nagai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.F.,
Goldman,G.H., Carvalho,A.F., Matsukuma,A., Baia,G.S., Simpson,D.H.,
Brunstein,A., deOliveira,P.S., Bucher,P., Jongeneel,C.V., O'Hare
,M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and

TITLE

JOURNAL
MEDLINE
COMMENT

Simpson,A.J.
Shotgun sequencing of the human transcriptome with ORF expressed
sequence tags
Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
20202663
Contact: Simpson A.J.G.
Laboratory of Cancer Genetics
Ludwig Institute for Cancer Research
Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
Brazil
Tel: +55-11-2704922
Fax: +55-11-2707001
Email: asimpson@ludwig.org.br
This sequence was derived from the FAPESP/LICR Human Cancer Genome
Project. This entry can be seen in the following URL
(http://www.ludwig.org.br/scripts/gethtml2.pl?t1=&t2=CM4-HT0744-160
600-201-a10&t3=2000-06-16&t4=1)
Seq primer: puc 18 forward
High quality sequence start: 11
High quality sequence stop: 370.

FEATURES
Location/Qualifiers

1. .370
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_lib="HT0744"
/dev_stage="Adult"
/note="Organ: head neck; Vector: puc18; Site_1: SmaI;
Site_2: SmaI; A mini-library was made by cloning products
derived from ORESTES PCR (U.S. Letters Patent application
No. 196,716 - Ludwig Institute for Cancer Research)
profiles into the pUC 18 vector. Reverse transcription of
tissue mRNA and cDNA amplification were performed under
low stringency conditions."
BASE COUNT 85 a 97 c 99 g 89 t
ORIGIN

Alignment Scores:
Pred. No.: 2.66e-29 Length: 370
Score: 327.00 Matches: 57
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 12 Gaps: 0

09-833799-13B (1-57) x BE715605 (1-370)

Qy 1 AlaGlnGluProValLysGlyProValSerThrLysProGlySerCysProIleLeu 20
Db 240 GCGCAAGAGCCAGTCAAAGTCCAGTCTCCACTAAGCGTGGCTCCTGCCCATATCTTG 181
Qy 21 IleArgCysAlaMetLeuAsnProProAsnArgCysLeuLysAspThrAspCysProGly 40
Db 180 ATCCGGTGGCCCATGTTGAATCCCCCTAACCGCTGCTTGAAGAGATACCTGACTGCCCAGGA 121
Qy 41 IleLysLysCysCysGluGlySerCysGlyMetAlaCysPheValProGln 57
Db 120 ATCAAGAAGTGCTGTGAAGGCTCTTGGGGATGGCCTGTTTCGTTCCCCAG 70

RESULT 10
BE715626

LOCUS BE715626 370 bp mRNA linear EST 12-SEP-2000
DEFINITION CM4-HT0744-160600-201-h09 HT0744 Homo sapiens cDNA, mRNA sequence.
ACCESSION BE715626
VERSION BE715626.1 GI:10103891
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 370)
AUTHORS Dias Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Briones,M.R.,
Nagai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.F.,
Goldman,G.H., Carvalho,A.F., Matsukuma,A., Baia,G.S., Simpson,D.H.,

TITLE
JOURNAL
MEDLINE
COMMENT

Brunstein,A., deOliveira,P.S., Bucher,P., Jongeneel,C.V., O'Hare
,M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and
Simpson,A.J.
Shotgun sequencing of the human transcriptome with ORF expressed
sequence tags
Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
20202663
Contact: Simpson A.J.G.
Laboratory of Cancer Genetics
Ludwig Institute for Cancer Research
Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
Brazil
Tel: +55-11-2704922
Fax: +55-11-2707001
Email: asimpson@ludwig.org.br
This sequence was derived from the FAPESP/LICR Human Cancer Genome
Project. This entry can be seen in the following URL
(http://www.ludwig.org.br/scripts/gethtml2.pl?tl=&t2=CM4-HT0744-160
600-201-h09&t3=2000-06-16&t4=1)
Seq primer: puc 18 forward
High quality sequence start: 13
High quality sequence stop: 370.

FEATURES
source

1. .370
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_lib="HT0744"
/dev_stage="Adult"
/note="Organ: head neck; Vector: puc18; Site_1: SmaI;
Site_2: SmaI; A mini-library was made by cloning products
derived from ORESTES PCR (U.S. Letters Patent application
No. 196,716 - Ludwig Institute for Cancer Research)
profiles into the pUC 18 vector. Reverse transcription of
tissue mRNA and cDNA amplification were performed under
low stringency conditions."
BASE COUNT 85 a 110 c 87 g 88 t
ORIGIN

Alignment Scores:
Pred. No.: 2.66e-29 Length: 370
Score: 327.00 Matches: 57
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 12 Gaps: 0

09-833799-13B (1-57) x BE715626 (1-370)

QY 1 AlaGlnGluProValLysGlyProValSerThrLysProGlySerCysProIleLeu 20
|||||
Db 97 GCGCAAGAGCCAGTCAAAGGTCCAGTCTCCACTAAGCTGGCTCTGCCCATATCTTG 156
|||||
QY 21 IleArgCysAlaMetLeuAsnProProAsnArgCysLeuLysAspThrAspCysProGly 40
|||||
Db 157 ATCCGGTGGCCCATGTTGAATCCCCCTAACCGCTGCTTGAAAGATACTGACTGCCCAGGA 216
|||||
QY 41 IleLysLysCysCysGluGlySerCysGlyMetAlaCysPheValProGln 57
|||||
Db 217 ATCAAGAAGTGCTGTGAAGGCTCTTGCGGGATGGCCTGTTTCGTTCCCCAG 267
|||||

RESULT 11
BE772160/c
LOCUS BE772160 370 bp mRNA linear EST 20-SEP-2000
DEFINITION CM4-FT0104-230600-215-c08 FT0104 Homo sapiens cDNA, mRNA sequence.
ACCESSION BE772160
VERSION BE772160.1 GI:10225818
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 370)
AUTHORS Dias Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Briones,M.R.,

TITLE
JOURNAL
MEDLINE
COMMENT

Nagai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.F.,
Goldman,G.H., Carvalho,A.F., Matsukuma,A., Baia,G.S., Simpson,D.H.,
Brunstein,A., deOliveira,P.S., Bucher,P., Jongeneel,C.V., O'Hare
,M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and
Simpson,A.J.
Shotgun sequencing of the human transcriptome with ORF expressed
sequence tags
Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
20202663
Contact: Simpson A.J.G.
Laboratory of Cancer Genetics
Ludwig Institute for Cancer Research
Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
Brazil
Tel: +55-11-2704922
Fax: +55-11-2707001
Email: asimpson@ludwig.org.br
This sequence was derived from the FAPESP/LICR Human Cancer Genome
Project. This entry can be seen in the following URL
(http://www.ludwig.org.br/scripts/gethtml2.pl?tl=&t2=CM4-FT0104-230
600-215-c08&t3=2000-06-23&t4=1)
Seq primer: puc 18 forward
High quality sequence start: 8
High quality sequence stop: 370.

FEATURES
source

1. .370
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_lib="FT0104"
/dev_stage="Adult"
/note="Organ: prostate tumor; Vector: puc18; Site_1: SmaI;
Site_2: SmaI; A mini-library was made by cloning products
derived from ORESTES PCR (U.S. Letters Patent application
No. 196,716 - Ludwig Institute for Cancer Research)
profiles into the pUC 18 vector. Reverse transcription of
tissue mRNA and cDNA amplification were performed under
low stringency conditions."
BASE COUNT 86 a 95 c 99 g 90 t
ORIGIN

Alignment Scores:
Pred. No.: 2.66e-29 Length: 370
Score: 327.00 Matches: 57
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 12 Gaps: 0

09-833799-13B (1-57) x BE772160 (1-370)

QY 1 AlaGlnGluProValLysGlyProValSerThrLysProGlySerCysProIleLeu 20
|||||
Db 239 GCGCAAGAGCCAGTCAAAGGTCCAGTCTCCACTAAGCTGGCTCTGCCCATATCTTG 180
|||||
QY 21 IleArgCysAlaMetLeuAsnProProAsnArgCysLeuLysAspThrAspCysProGly 40
|||||
Db 179 ATCCGGTGGCCCATGTTGAATCCCCCTAACCGCTGCTTGAAAGATACTGACTGCCCAGGA 120
|||||
QY 41 IleLysLysCysCysGluGlySerCysGlyMetAlaCysPheValProGln 57
|||||
Db 119 ATCAAGAAGTGCTGTGAAGGCTCTTGCGGGATGGCCTGTTTCGTTCCCCAG 69
|||||

RESULT 12
BF838555
LOCUS BF838555 371 bp mRNA linear EST 13-JAN-2001
DEFINITION IL5-HT0990-251100-263-g08 HT0990 Homo sapiens cDNA, mRNA sequence.
ACCESSION BF838555
VERSION BF838555.1 GI:12190853
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE 1 (bases 1 to 371)
AUTHORS Dias Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Briones,M.R., Nagai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.F., Goldman,G.H., Carvalho,A.F., Matsukuma,A., Baia,G.S., Simpson,D.H., Brunstein,A., deOliveira,P.S., Bucher,P., Jongeneel,C.V., O'Hare ,M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and Simpson,A.J.
TITLE Shotgun sequencing of the human transcriptome with ORF expressed sequence tags
JOURNAL Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
MEDLINE 20202663
COMMENT Contact: Simpson A.J.G.
Laboratory of Cancer Genetics
Ludwig Institute for Cancer Research
Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP, Brazil
Tel: +55-11-2704922
Fax: +55-11-2707001
Email: asimpson@ludwig.org.br
This sequence was derived from the FAPESP/LICR Human Cancer Genome Project. This entry can be seen in the following URL
(http://www.ludwig.org.br/scripts/gethtml2.pl?tl=IL5&t2=IL5-HT0990-251100-263-g08&t3=2000-11-25&t4=1)
Seq primer: puc 18 forward
High quality sequence stop: 367.

FEATURES
source
1..371
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_lib="HT0990"
/dev_stage="Adult"
/note="Organ: head neck; Vector: puc18; Site 1: SmaI;
Site_2: SmaI; A mini-library was made by cloning products
derived from ORESTES PCR (U.S. Letters Patent application
No. 196,716 - Ludwig Institute for Cancer Research)
profiles into the pUC 18 vector. Reverse transcription of
tissue mRNA and cDNA amplification were performed under
low stringency conditions."
BASE COUNT 82 a 117 c 85 g 87 t
ORIGIN

Alignment Scores:
Pred. No.: 2.67e-29 Length: 371
Score: 327.00 Matches: 57
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 12 Gaps: 0

09-833799-13B (1-57) x BF838555 (1-371)
QY 1 AlaGlnGluProValLysGlyProValSerThrLysProGlySerCysProIleLeu 20
|||||
Db 77 GCGCAAGAGCCAGTCAAAGGTCAGTCTCCACTAAGCTGGCTCCTGCCCATATCTTG 136
|||||
QY 21 IleArgCysAlaMetLeuAsnProProAsnArgCysLeuLysAspThrAspCysProGly 40
|||||
Db 137 ATCCGGTGGCCCATGTTGATCCCCCTAACCGCTGCTTGAAAGATACGTGACTGCCCAGGA 196
|||||
QY 41 IleLysLysCysCysGluGlySerCysGlyMetAlaCysPheValProGln 57
|||||
Db 197 ATCAAGAAGTGCTGTGAAGGCTCTTGCGGGATGGCCTGTTTCGTTCCCCAG 247
|||||

RESULT 13
BE772166/c
LOCUS BE772166 371 bp mRNA linear EST 20-SEP-2000
DEFINITION CM4-FT0104-230600-215-d05 FT0104 Homo sapiens cDNA, mRNA sequence.
ACCESSION BE772166
VERSION BE772166.1 GI:10225824
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 371)
Dias Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Briones,M.R., Nagai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.F., Goldman,G.H., Carvalho,A.F., Matsukuma,A., Baia,G.S., Simpson,D.H., Brunstein,A., deOliveira,P.S., Bucher,P., Jongeneel,C.V., O'Hare ,M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and Simpson,A.J.
TITLE Shotgun sequencing of the human transcriptome with ORF expressed sequence tags
JOURNAL Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
MEDLINE 20202663
COMMENT Contact: Simpson A.J.G.
Laboratory of Cancer Genetics
Ludwig Institute for Cancer Research
Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP, Brazil
Tel: +55-11-2704922
Fax: +55-11-2707001
Email: asimpson@ludwig.org.br
This sequence was derived from the FAPESP/LICR Human Cancer Genome Project. This entry can be seen in the following URL
(http://www.ludwig.org.br/scripts/gethtml2.pl?tl=CM4-FT0104-230600-215-d05&t3=2000-06-23&t4=1)
Seq primer: puc 18 forward
High quality sequence stop: 371.
High quality sequence start: 371.

FEATURES
source
1..371
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_lib="FT0104"
/dev_stage="Adult"
/note="Organ: prostate tumor; Vector: puc18; Site 1: SmaI;
Site_2: SmaI; A mini-library was made by cloning products
derived from ORESTES PCR (U.S. Letters Patent application
No. 196,716 - Ludwig Institute for Cancer Research)
profiles into the pUC 18 vector. Reverse transcription of
tissue mRNA and cDNA amplification were performed under
low stringency conditions."
BASE COUNT 84 a 96 c 101 g 90 t
ORIGIN

Alignment Scores:
Pred. No.: 2.67e-29 Length: 371
Score: 327.00 Matches: 57
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 12 Gaps: 0

09-833799-13B (1-57) x BE772166 (1-371)
QY 1 AlaGlnGluProValLysGlyProValSerThrLysProGlySerCysProIleLeu 20
|||||
Db 240 GCGCAAGAGCCAGTCAAAGGTCAGTCTCCACTAAGCTGGCTCCTGCCCATATCTTG 181
|||||
QY 21 IleArgCysAlaMetLeuAsnProProAsnArgCysLeuLysAspThrAspCysProGly 40
|||||
Db 180 ATCCGGTGGCCCATGTTGATCCCCCTAACCGCTGCTTGAAAGATACGTGACTGCCCAGGA 121
|||||
QY 41 IleLysLysCysCysGluGlySerCysGlyMetAlaCysPheValProGln 57
|||||
Db 120 ATCAAGAAGTGCTGTGAAGGCTCTTGCGGGATGGCCTGTTTCGTTCCCCAG 70
|||||

RESULT 14
AI285400/c
LOCUS AI285400 375 bp mRNA linear EST 23-NOV-1998
DEFINITION qt69f06.x1 NCI CGAP Eso2 Homo sapiens cDNA clone IMAGE:1960547 3', similar to gb:SS8717_rna1 ELAFIN PRECURSOR (HUMAN); contains element THR repetitive element ;, mRNA sequence.
ACCESSION AI285400
VERSION AI285400.1 GI:3923633

KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 375)
AUTHORS NCI-CGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.
TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
JOURNAL Unpublished (1997)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgaps-r@mail.nih.gov
Tissue Procurement: Nan Hu, M.D., Ph.D., Mark Roth, M.D., Phillip
Taylor, M.D., Michael R. Emmert-Buck, M.D., Ph.D.
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: Greg Lennon, Ph.D.
DNA Sequencing by: Washington University Genome Sequencing Center
Clone distribution: NCI-CGAP clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
www-bio.llnl.gov/bbrp/image/image.html
Seq primer: -40UP from Gibco
High quality sequence stop: 356.
Location/Qualifiers
1. 375

FEATURES
source
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:1960547"
/clone_lib="NCI_CGAP_Eso2"
/tissue_type="squamous cell carcinoma"
/lab_host="DH10B"
/note="Organ: esophagus; Vector: pCMV-SPORT6; Site 1: SalI
; Site 2: NotI; Cloned unidirectionally. Primer: Oligo
dr. Average insert size 1.1 kb. Life Technologies catalog
#: 11502-010"
BASE COUNT 91 a 83 c 125 g 75 t 1 others
ORIGIN

Alignment Scores:
Pred. No.: 2.71e-29 Length: 375
Score: 327.00 Matches: 57
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 9 Gaps: 0

09-833799-13B (1-57) x AI285400 (1-375)
QY 1 AlaGlnGluProValLysGlyProValSerThrLysProGlySerCysProIleLeu 20
|||||
Db 349 GCGAAGAGCCAGTCAAAGGTCCAGTCTCCACTAAGCCTGGCTCCTGCCCATTTCTTG 290
QY 21 IleArgCysAlaMetLeuAsnProProAsnArgCysLeuLysAspThrAspCysProGly 40
|||||
Db 289 ATCCGGTGCGCCATGTTGAATCCCCCTAACCGCTGCTTGAAGATACTGACTGCCCAGGA 230
QY 41 IleLysLysCysCysGluGlySerCysGlyMetAlaCysPheValProGln 57
|||||
Db 229 ATCAAGAAGTCTGTGAAGGCTCTTGCAGGATGGCCTGTTTCGTTCCCCAG 179
RESULT 15
BE715603/c 379 bp mRNA linear EST 12-SEP-2000
LOCUS CM4-HT0744-160600-201-a07 HT0744 Homo sapiens cDNA, mRNA sequence.
DEFINITION BE715603
ACCESSION BE715603
VERSION BE715603.1 GI:10103868
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 379)
AUTHORS Dias Neto, E., Garcia Correa, R., Verjovski-Almeida, S., Briones, M.R.,
Nagai, M.A., da Silva, W. Jr., Zago, M.A., Bordin, S., Costa, F.F.,

Goldman, G.H., Carvalho, A.F., Matsukuma, A., Baia, G.S., Simpson, D.H.,
Brunstein, A., deOliveira, P.S., Bucher, P., Jongeneel, C.V., O'Hare
M.J., Soares, F., Brentani, R.R., Reis, L.F., de Souza, S.J. and
Simpson, A.J.
Shotgun sequencing of the human transcriptome with ORF expressed
sequence tags
Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
20202663
Contact: Simpson A.J.G.
Laboratory of Cancer Genetics
Ludwig Institute for Cancer Research
Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
Brazil
Tel: +55-11-2704922
Fax: +55-11-2707001
Email: asimpson@ludwig.org.br
This sequence was derived from the FAPESP/LICR Human Cancer Genome
Project. This entry can be seen in the following URL
(<http://www.ludwig.org.br/scripts/gethtml2.pl?t1=&t2=CM4-HT0744-160>)
600-201-a07&t3=2000-06-16&t4=1)
Seq primer: puc 18 forward
High quality sequence start: 25
High quality sequence stop: 379.
Location/Qualifiers
1. 379

FEATURES
source
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_lib="HT0744"
/dev_stage="Adult"
/note="Organ: head neck; Vector: puc18; Site 1: SmaI;
Site 2: SmaI; A mini-library was made by cloning products
derived from ORESTES PCR (U.S. Letters Patent application
No. 196,716 - Ludwig Institute for Cancer Research)
profiles into the pUC 18 vector. Reverse transcription of
tissue mRNA and cDNA amplification were performed under
low stringency conditions."
BASE COUNT 86 a 95 c 112 g 86 t
ORIGIN

Alignment Scores:
Pred. No.: 2.74e-29 Length: 379
Score: 327.00 Matches: 57
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 12 Gaps: 0

09-833799-13B (1-57) x BE715603 (1-379)
QY 1 AlaGlnGluProValLysGlyProValSerThrLysProGlySerCysProIleLeu 20
|||||
Db 260 GCGAAGAGCCAGTCAAAGGTCCAGTCTCCACTAAGCCTGGCTCCTGCCCATTTCTTG 201
QY 21 IleArgCysAlaMetLeuAsnProProAsnArgCysLeuLysAspThrAspCysProGly 40
|||||
Db 200 ATCCGGTGCGCCATGTTGAATCCCCCTAACCGCTGCTTGAAGATACTGACTGCCCAGGA 141
QY 41 IleLysLysCysCysGluGlySerCysGlyMetAlaCysPheValProGln 57
|||||
Db 140 ATCAAGAAGTCTGTGAAGGCTCTTGCAGGATGGCCTGTTTCGTTCCCCAG 90
RESULT 16
BE772155/c 380 bp mRNA linear EST 20-SEP-2000
LOCUS BE772155
DEFINITION CM4-FT0104-230600-215-b05 FT0104 Homo sapiens cDNA, mRNA sequence.
ACCESSION BE772155
VERSION BE772155.1 GI:10225813
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 380)

AUTHORS Dias Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Briones,M.R., Nagai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.F., Goldman,G.H., Carvalho,A.F., Matsukuma,A., Baia,G.S., Simpson,D.H., Brunstein,A., deOliveira,P.S., Bucher,P., Jongeneel,C.V., O'Hare ,M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and Simpson,A.J.

TITLE Shotgun sequencing of the human transcriptome with ORF expressed sequence tags

JOURNAL Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)

MEDLINE 20202663

COMMENT Contact: Simpson A.J.G.
Laboratory of Cancer Genetics
Ludwig Institute for Cancer Research
Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP, Brazil
Tel: +55-11-2704922
Fax: +55-11-2707001
Email: asimpson@ludwig.org.br
This sequence was derived from the FAPESP/LICR Human Cancer Genome Project. This entry can be seen in the following URL
(http://www.ludwig.org.br/scripts/gethtml2.pl?tl=&t3=CM4-FT0104-230 600-215-b05&t3=2000-06-23&t4=1)
Seq primer: puc 18 forward
High quality sequence start: 10
High quality sequence stop: 365.

FEATURES
source
1. .380
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_lib="FT0104"
/dev_stage="Adult"
/note="Organ: prostate_tumor; Vector: puc18; Site_1: Smal;
Site_2: Smal; A mini-library was made by cloning products
derived from ORESTES PCR (U.S. Letters Patent application
No. 196,716 - Ludwig Institute for Cancer Research)
profiles into the pUC 18 vector. Reverse transcription of
tissue mRNA and cDNA amplification were performed under
low stringency conditions."
89 a 99 c 100 g 92 t

BASE COUNT 89 a 99 c 100 g 92 t

ORIGIN

Alignment Scores:
Pred. No.: 2.75e-29 Length: 380
Score: 327.00 Matches: 57
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 12 Gaps: 0

09-833799-13B (1-57) x BE772155 (1-380)

QY 1 AlaGlnGluProValLysGlyProValSerThrLysProGlySerCysProIleLeu 20
|||||
Db 247 GCGCAAGAGCCAGTCAAAGGTCAGTCTCCACTAAGCCTGGCTCCTGCCCATTAATCTTG 188
|||||
QY 21 IleArgCysAlaMetLeuAsnProProAsnArgCysLeuLysAspThrAspCysProGly 40
|||||
Db 187 ATCCGGTGCCTCATGTTGAATCCCTTAACCGCTGCTTGAAAGATATACTGACTGCCCAGGA 128
|||||
QY 41 IleLysLysCysCysGluGlySerCysGlyMetAlaCysPheValProGln 57
|||||
Db 127 ATCAAGAAGTGTGTGAAGGCTCTTGGGGATGGCCTGTTTCGTTCCCCAG 77
|||||

RESULT 17
BE772171/c
LOCUS BE772171 380 bp mRNA linear EST 20-SEP-2000
DEFINITION CM4-FT0104-230600-215-g07 FT0104 Homo sapiens cDNA, mRNA sequence.
ACCESSION BE772171
VERSION BE772171.1 GI:10225829
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

REFERENCE 1 (bases 1 to 380)
AUTHORS Dias Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Briones,M.R., Nagai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.F., Goldman,G.H., Carvalho,A.F., Matsukuma,A., Baia,G.S., Simpson,D.H., Brunstein,A., deOliveira,P.S., Bucher,P., Jongeneel,C.V., O'Hare ,M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and Simpson,A.J.

TITLE Shotgun sequencing of the human transcriptome with ORF expressed sequence tags

JOURNAL Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)

MEDLINE 20202663

COMMENT Contact: Simpson A.J.G.
Laboratory of Cancer Genetics
Ludwig Institute for Cancer Research
Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP, Brazil
Tel: +55-11-2704922
Fax: +55-11-2707001
Email: asimpson@ludwig.org.br
This sequence was derived from the FAPESP/LICR Human Cancer Genome Project. This entry can be seen in the following URL
(http://www.ludwig.org.br/scripts/gethtml2.pl?tl=&t3=CM4-FT0104-230 600-215-g07&t3=2000-06-23&t4=1)
Seq primer: puc 18 forward
High quality sequence start: 23
High quality sequence stop: 380.

FEATURES
source
1. .380
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_lib="FT0104"
/dev_stage="Adult"
/note="Organ: prostate_tumor; Vector: puc18; Site_1: Smal;
Site_2: Smal; A mini-library was made by cloning products
derived from ORESTES PCR (U.S. Letters Patent application
No. 196,716 - Ludwig Institute for Cancer Research)
profiles into the pUC 18 vector. Reverse transcription of
tissue mRNA and cDNA amplification were performed under
low stringency conditions."
89 a 97 c 106 g 88 t

BASE COUNT 89 a 97 c 106 g 88 t

ORIGIN

Alignment Scores:
Pred. No.: 2.75e-29 Length: 380
Score: 327.00 Matches: 57
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 12 Gaps: 0

09-833799-13B (1-57) x BE772171 (1-380)

QY 1 AlaGlnGluProValLysGlyProValSerThrLysProGlySerCysProIleLeu 20
|||||
Db 251 GCGCAAGAGCCAGTCAAAGGTCAGTCTCCACTAAGCCTGGCTCCTGCCCATTAATCTTG 192
|||||
QY 21 IleArgCysAlaMetLeuAsnProProAsnArgCysLeuLysAspThrAspCysProGly 40
|||||
Db 191 ATCCGGTGCCTCATGTTGAATCCCTTAACCGCTGCTTGAAAGATATACTGACTGCCCAGGA 132
|||||
QY 41 IleLysLysCysCysGluGlySerCysGlyMetAlaCysPheValProGln 57
|||||
Db 131 ATCAAGAAGTGTGTGAAGGCTCTTGGGGATGGCCTGTTTCGTTCCCCAG 81
|||||

RESULT 18
BE182879
LOCUS BE182879 386 bp mRNA linear EST 22-JUN-2000
DEFINITION CM4-HT0652-150400-143-g08 HT0652 Homo sapiens cDNA, mRNA sequence.
ACCESSION BE182879
VERSION BE182879.1 GI:8662055
KEYWORDS EST.
SOURCE human.

ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 386)
Dias Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Briones,M.R.,
Nagai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.F.,
Goldman,G.H., Carvalho,A.F., Matsukuma,A., Baia,G.S., Simpson,D.H.,
Brunstein,A., deOliveira,P.S., Bucher,P., Jongeneel,C.V., O'Hare
M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and
Simpson,A.J.
Shotgun sequencing of the human transcriptome with ORF expressed
sequence tags
Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
20202663
Contact: Simpson A.J.G.
Laboratory of Cancer Genetics
Ludwig Institute for Cancer Research
Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
Brazil
Tel: +55-11-2704922
Fax: +55-11-2707001
Email: asimpson@ludwig.org.br
This sequence was derived from the FAPESP/LICR Human Cancer Genome
Project. This entry can be seen in the following URL
(http://www.ludwig.org.br/scripts/gethtml2.pl?tl=&t2=CM4-HT0652-150
400-143-g08&t3=2000-04-15&t4=1)
Seq primer: puc 18 forward
High quality sequence stop: 61.
Location/Qualifiers
1. .386
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_lib="HT0652"
/dev_stage="Adult"
/note="Organ: head_neck; Vector: puc18; Site_1: SmaI;
Site_2: SmaI; A mini-library was made by cloning products
derived from ORESTES PCR (U.S. Letters Patent application
No. 196,716 - Ludwig Institute for Cancer Research)
profiles into the pUC 18 vector. Reverse transcription of
tissue mRNA and cDNA amplification were performed under
low stringency conditions."
89 a 104 c 105 g 88 t
BASE COUNT
ORIGIN
Alignment Scores:
Pred. No.: 2.8e-29 Length: 386
Score: 327.00 Matches: 57
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 10 Gaps: 0
09-833799-13B (1-57) x BE182879 (1-386)
QY 1 AlaGlnGluProValLysGlyProValSerThrLysProGlySerCysProIleLeu 20
Db 146 GCGAAGAGCCAGTCAAAGGTCCAGTCTCCACTAAGCTGGCTCCTGCCCATATATCTTG 205
QY 21 IleArgCysAlaMetLeuAsnProProAsnArgCysLeuLysAspThrAspCysProGly 40
Db 206 ATCCGGTGGCCCATGTTGAATCCCCCTAACCGCTGCTGAAAGATACTGACTGCCCAGGA 265
QY 41 IleLysLysCysCysGluGlySerCysGlyMetAlaCysPheValProGln 57
Db 266 ATCAAGAAGTGCTGTGAAGGCTCTTGGGGATGGCTGTTTCGTTCCCCAG 316
RESULT 19
BE772169/c
LOCUS BE772169 386 bp mRNA linear EST 20-SEP-2000
DEFINITION CM4-FT0104-230600-215-f02 FT0104 Homo sapiens cDNA, mRNA sequence.
ACCESSION BE772169
VERSION BE772169.1 GI:10225827
KEYWORDS EST.

SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 386)
Dias Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Briones,M.R.,
Nagai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.F.,
Goldman,G.H., Carvalho,A.F., Matsukuma,A., Baia,G.S., Simpson,D.H.,
Brunstein,A., deOliveira,P.S., Bucher,P., Jongeneel,C.V., O'Hare
M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and
Simpson,A.J.
Shotgun sequencing of the human transcriptome with ORF expressed
sequence tags
Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
20202663
Contact: Simpson A.J.G.
Laboratory of Cancer Genetics
Ludwig Institute for Cancer Research
Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
Brazil
Tel: +55-11-2704922
Fax: +55-11-2707001
Email: asimpson@ludwig.org.br
This sequence was derived from the FAPESP/LICR Human Cancer Genome
Project. This entry can be seen in the following URL
(http://www.ludwig.org.br/scripts/gethtml2.pl?tl=&t2=CM4-FT0104-230
600-215-f02&t3=2000-06-23&t4=1)
Seq primer: puc 18 forward
High quality sequence start: 45
High quality sequence stop: 386.
Location/Qualifiers
1. .386
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_lib="FT0104"
/dev_stage="Adult"
/note="Organ: prostate tumor; Vector: puc18; Site_1: SmaI;
Site_2: SmaI; A mini-library was made by cloning products
derived from ORESTES PCR (U.S. Letters Patent application
No. 196,716 - Ludwig Institute for Cancer Research)
profiles into the pUC 18 vector. Reverse transcription of
tissue mRNA and cDNA amplification were performed under
low stringency conditions."
90 a 99 c 107 g 90 t
BASE COUNT
ORIGIN
Alignment Scores:
Pred. No.: 2.8e-29 Length: 386
Score: 327.00 Matches: 57
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 12 Gaps: 0
09-833799-13B (1-57) x BE772169 (1-386)
QY 1 AlaGlnGluProValLysGlyProValSerThrLysProGlySerCysProIleLeu 20
Db 255 GCGAAGAGCCAGTCAAAGGTCCAGTCTCCACTAAGCTGGCTCCTGCCCATATATCTTG 196
QY 21 IleArgCysAlaMetLeuAsnProProAsnArgCysLeuLysAspThrAspCysProGly 40
Db 195 ATCCGGTGGCCCATGTTGAATCCCCCTAACCGCTGCTGAAAGATACTGACTGCCCAGGA 136
QY 41 IleLysLysCysCysGluGlySerCysGlyMetAlaCysPheValProGln 57
Db 135 ATCAAGAAGTGCTGTGAAGGCTCTTGGGGATGGCTGTTTCGTTCCCCAG 85
RESULT 20
AA583997/c
LOCUS AA583997 388 bp mRNA linear EST 26-SEP-1997
DEFINITION nm5a01.s1 NCI CGAP Lar1 Homo sapiens cDNA clone IMAGE:1088712 3'
similar to gb:S58717_rnal ELAFIN PRECURSOR (HUMAN);, mRNA sequence.

ACCESSION AA583997
VERSION AA583997.1 GI:2368606
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 388)
AUTHORS NCI-CGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.
TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
JOURNAL Unpublished (1997)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: L. Jeffrey Medeiros, M.D., Michael R.
Emmert-Buck, M.D., Ph.D.
cDNA Library Preparation: Stratagene, Inc.
CDNA Library Arrayed by: Greg Lennon, Ph.D.
DNA Sequencing by: Washington University Genome Sequencing Center
Clone distribution: NCI-CGAP clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
www-bio.llnl.gov/bbrp/image/image.html
Insert Length: 480 Std Error: 0.00
Seq primer: -40ml3 fwd. ET from Amersham.
FEATURES
source
1..388
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_image="1088712"
/clone_lib="NCI-CGAP_Lar1"
/tissue_type="larynx"
/lab_host="SOLR (kanamycin resistant)"
/note="Organ: larynx; Vector: Bluescript SK-; Site_1:
EcoRI; Site_2: XhoI; Cloned unidirectionally. Primer:
Oligo dT. Larynx. 5' adaptor sequence: 5' GAATTCGGCAGCAG
3' 3' adaptor sequence: 5' CTCGAGTTTTTTTTTTTTTTT 3'
Average insert size: 0.9 kb."
BASE COUNT 91 a 83 c 124 g 90 t
ORIGIN
Alignment Scores:
Pred. No.: 2.82e-29 Length: 388
Score: 327.00 Matches: 57
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 9 Gaps: 0
09-833799-13B (1-57) x AA583997 (1-388)
QY 1 AlaGlnGluProValLysGlyProValSerThrLysProGlySerCysProIleLeu 20
|||||
Db 363 GCGCAAGAGCCAGTCAAAGGTCAGTCTCCACTAAGCGTGCTCTGCCCATATCTTG 304
|||||
QY 21 IleArgCysAlaMetLeuAsnProProAsnArgCysLeuLysAspThrAspCysProGly 40
|||||
Db 303 ATCCGGTGGCCATGTTGAATCCCCCTAACCGCTGCTTGAAAGATACCTGACTGCCAGGA 244
|||||
QY 41 IleLysLysCysCysGluGlySerCysGlyMetAlaCysPheValProGln 57
|||||
Db 243 ATCAAGAAGTGCTGTGAAGGCTCTTGCGGGATGCGCTGTTTCGTTCCCCAG 193
|||||
RESULT 21
BE772159
LOCUS BE772159 388 bp mRNA linear EST 20-SEP-2000
DEFINITION CM4-FT0104-230600-215-c06 FT0104 Homo sapiens cDNA, mRNA sequence.
ACCESSION BE772159
VERSION BE772159.1 GI:10225817
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE 1 (bases 1 to 388)
AUTHORS Dias Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Briones,M.R.,
Nagai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.F.,
Goldman,G.H., Carvalho,A.F., Matsukuma,A., Baia,G.S., Simpson,D.H.,
Brunstein,A., deOliveira,P.S., Bucher,P., Jongeneel,C.V., O'Hare
M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and
Simpson,A.J.
TITLE Shotgun sequencing of the human transcriptome with ORF expressed
sequence tags
JOURNAL Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
MEDLINE 20202663
COMMENT Contact: Simpson A.J.G.
Laboratory of Cancer Genetics
Ludwig Institute for Cancer Research
Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
Brazil
Tel: +55-11-2704922
Fax: +55-11-2707001
Email: asimpson@ludwig.org.br
This sequence was derived from the FAPESP/LICR Human Cancer Genome
Project. This entry can be seen in the following URL
(<http://www.ludwig.org.br/scripts/gethtml2.pl?tl=&t2=CM4-FT0104-230600-215-c06&t3=2000-06-23&t4=1>)
Seq primer: puc 18 forward
High quality sequence start: 26
High quality sequence stop: 388.
FEATURES
source
1..388
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_lib="FT0104"
/dev_stage="Adult"
/note="Organ: prostate tumor; Vector: puc18; Site_1: SmaI;
Site_2: SmaI; A mini-library was made by cloning products
derived from ORESTES PCR (U.S. Letters Patent application
No. 196,716 - Ludwig Institute for Cancer Research)
profiles into the pUC 18 vector. Reverse transcription of
tissue mRNA and cDNA amplification were performed under
low stringency conditions."
BASE COUNT 88 a 115 c 93 g 92 t
ORIGIN
Alignment Scores:
Pred. No.: 2.82e-29 Length: 388
Score: 327.00 Matches: 57
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 12 Gaps: 0
09-833799-13B (1-57) x BE772159 (1-388)
QY 1 AlaGlnGluProValLysGlyProValSerThrLysProGlySerCysProIleLeu 20
|||||
Db 115 GCGCAAGAGCCAGTCAAAGGTCAGTCTCCACTAAGCGTGCTCTGCCCATATCTTG 174
|||||
QY 21 IleArgCysAlaMetLeuAsnProProAsnArgCysLeuLysAspThrAspCysProGly 40
|||||
Db 175 ATCCGGTGGCCATGTTGAATCCCCCTAACCGCTGCTTGAAAGATACCTGACTGCCAGGA 234
|||||
QY 41 IleLysLysCysCysGluGlySerCysGlyMetAlaCysPheValProGln 57
|||||
Db 235 ATCAAGAAGTGCTGTGAAGGCTCTTGCGGGATGCGCTGTTTCGTTCCCCAG 285
|||||
RESULT 22
BE772168
LOCUS BE772168 389 bp mRNA linear EST 20-SEP-2000
DEFINITION CM4-FT0104-230600-215-d07 FT0104 Homo sapiens cDNA, mRNA sequence.
ACCESSION BE772168
VERSION BE772168.1 GI:10225826
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens

VERSION AI989737.1 GI:5836618
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 390)
AUTHORS NCI-CGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.
TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
JOURNAL Unpublished (1997)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
This clone is available royalty-free through LLNL ; contact the
IMAGE Consortium (info@image.llnl.gov) for further information.
Seq primer: -40Up from Gibco.
FEATURES
source
1..390
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:2520710"
/clone_lib="Soares_Dieckgraefe_colon_NHCD"
/tissue_type="colonic mucosa from 3 patients with Crohn's
disease"
/lab_host="DH10B (phage-resistant)"
/note="Organ: colon; Vector: pT7T3D-Pac (Pharmacia) with a
modified polylinker; Site 1: Not I; Site 2: Eco RI; 1st
strand cDNA was primed with a Not I - oligo(dT) primer [5'
TGTTACCAATCTGAAGTGGAGCGCGCGCTCTTTTCTTTTCTTTT 3'],
double-stranded cDNA was ligated to Eco RI adaptors
(Pharmacia), digested with Not I and cloned into the Not I
and Eco RI sites of the modified pT7T3 vector. Library
went through one round of normalization. Tissue samples
provided by Dr. Brian Dieckgraefe (Washington University,
dieck@im.wustl.edu); colonic mucosa represents a range of
disease involvement from moderate to severe Crohn's
disease; samples include both perforating (fistulas) and
non-perforating samples. Library constructed by Bento
Soares and M. Fatima Bonaldo."
BASE COUNT 94 a 86 c 126 g 84 t
ORIGIN
Alignment Scores:
Pred. No.: 2.84e-29 Length: 390
Score: 327.00 Matches: 57
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 9 Gaps: 0
09-833799-13B (1-57) x AI989737 (1-390)
Qy 1 AlaGlnGluProValLysGlyProValSerThrLysProGlySerCysProIleLeu 20
Db 354 GCGCAAGAGCCAGTCAAAGGTCAGTCTCCACTAAGCCTGGCTCCTGCCCATATCTTG 295
Qy 21 IleArgCysAlaMetLeuAsnProProAsnArgCysLeuLysAspThrAspCysProGly 40
Db 294 ATCCGGTGGCCCATGTTGAATCCCCCTAACCGCTGCTTGAAAGATACCTGACTGCCAGGA 235
Qy 41 IleLysLysCysCysGluGlySerCysGlyMetAlaCysPheValProGln 57
Db 234 ATCAAGAAAGTGTGTGAAGGCTCTTGCGGGATGGCCTGTTTCGTTCCCCAG 184
RESULT 25
BE715611/c 390 bp mRNA linear EST 12-SEP-2000
LOCUS BE715611
DEFINITION CM4-HT0744-160600-201-c07 HT0744 Homo sapiens cDNA, mRNA sequence.
ACCESSION BE715611
VERSION BE715611.1 GI:10103876
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 390)
AUTHORS Dias Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Briones,M.R.,
Nagai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.F.,
Goldman,G.H., Carvalho,A.F., Matsukuma,A., Baia,G.S., Simpson,D.H.,
Brunstein,A., deOliveira,P.S., Bucher,P., Jongeneel,C.V., O'Hare
M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and
Simpson,A.J.
TITLE Shotgun sequencing of the human transcriptome with ORF expressed
sequence tags
JOURNAL Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
MEDLINE 20202663
COMMENT Contact: Simpson A.J.G.
Laboratory of Cancer Genetics
Ludwig Institute for Cancer Research
Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
Brazil
Tel: +55-11-2704922
Fax: +55-11-2707001
Email: asimpson@ludwig.org.br
This sequence was derived from the FAPESP/LICR Human Cancer Genome
Project. This entry can be seen in the following URL
(<http://www.ludwig.org.br/scripts/gethtml2.pl?tl=&t2=CM4-HT0744-160>
600-201-c07&t3=2000-06-16&t4=1)
Seq primer: puc 18 forward
High quality sequence start: 30
High quality sequence stop: 390.
FEATURES
source
1..390
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_lib="HT0744"
/dev_stage="Adult"
/note="Organ: head neck; Vector: puc18; Site 1: SmaI;
Site 2: SmaI; A mini-library was made by cloning products
derived from ORESTES PCR (U.S. Letters Patent application
No. 196,716 - Ludwig Institute for Cancer Research)
profiles into the pUC 18 vector. Reverse transcription of
tissue mRNA and cDNA amplification were performed under
low stringency conditions."
BASE COUNT 90 a 97 c 111 g 92 t
ORIGIN
Alignment Scores:
Pred. No.: 2.84e-29 Length: 390
Score: 327.00 Matches: 57
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 12 Gaps: 0
09-833799-13B (1-57) x BE715611 (1-390)
Qy 1 AlaGlnGluProValLysGlyProValSerThrLysProGlySerCysProIleLeu 20
Db 259 GCGCAAGAGCCAGTCAAAGGTCAGTCTCCACTAAGCCTGGCTCCTGCCCATATCTTG 200
Qy 21 IleArgCysAlaMetLeuAsnProProAsnArgCysLeuLysAspThrAspCysProGly 40
Db 199 ATCCGGTGGCCCATGTTGAATCCCCCTAACCGCTGCTTGAAAGATACCTGACTGCCAGGA 140
Qy 41 IleLysLysCysCysGluGlySerCysGlyMetAlaCysPheValProGln 57
Db 139 ATCAAGAAAGTGTGTGAAGGCTCTTGCGGGATGGCCTGTTTCGTTCCCCAG 89
RESULT 26
BE715615 395 bp mRNA linear EST 12-SEP-2000
LOCUS BE715615
DEFINITION CM4-HT0744-160600-201-d10 HT0744 Homo sapiens cDNA, mRNA sequence.
ACCESSION BE715615
VERSION BE715615.1 GI:10103880
KEYWORDS EST.

SOURCE human.
ORGANISM Homo sapiens
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 395)
Dias Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Briones,M.R.,
Nagai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.F.,
Goldman,G.H., Carvalho,A.F., Matsukuma,A., Baia,G.S., Simpson,D.H.,
Brunstein,A., deOliveira,P.S., Bucher,P., Jongeneel,C.V., O'Hare
,M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and
Simpson,A.J.
TITLE Shotgun sequencing of the human transcriptome with ORF expressed
sequence tags
JOURNAL Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
MEDLINE 20202663
COMMENT Contact: Simpson A.J.G.
Laboratory of Cancer Genetics
Ludwig Institute for Cancer Research
Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
Brazil
Tel: +55-11-2704922
Fax: +55-11-2707001
Email: asimpson@ludwig.org.br
This sequence was derived from the FAPESP/LICR Human Cancer Genome
Project. This entry can be seen in the following URL
(http://www.ludwig.org.br/scripts/gethtml2.pl?t1=&t2=CM4-HT0744-160
600-201-dl0&t3=2000-06-16&t4=1)
Seq primer: puc 18 forward
High quality sequence start: 4
High quality sequence stop: 395.
Features Location/Qualifiers
1..395
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_lib="HT0744"
/dev_stage="Adult"
/note="Organ: head neck; Vector: puc18; Site_1: SmaI;
Site_2: SmaI; A mini-library was made by cloning products
derived from ORESTES PCR (U.S. Letters Patent application
No. 196,716 - Ludwig Institute for Cancer Research)
profiles into the pUC 18 vector. Reverse transcription of
tissue mRNA and cDNA amplification were performed under
low stringency conditions."
BASE COUNT 91 a 115 c 96 g 93 t
ORIGIN
Alignment Scores:
Pred. No.: 2.88e-29 Length: 395
Score: 327.00 Matches: 57
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 12 Gaps: 0
09-833799-13B (1-57) x BE715615 (1-395)
QY 1 AlaGlnGluProValLysGlyProValSerThrLysProGlySerCysProIleLeu 20
Db 122 GCGCAAGAGCCAGTCAAAGTCCAGTCTCCACTAAGCTGGCTCCTGCCCAATTATCTTG 181
QY 21 IleArgCysAlaMetLeuAsnProProAsnArgCysLeuLysAspThrAspCysProGly 40
Db 182 ATCCGGTGGCCCATGTTGAATCCCCCTTAACCGCTGCTTGAAGATATACTGACTGCCCAGGA 241
QY 41 IleLysLysCysCysGluGlySerCysGlyMetAlaCysPheValProGln 57
Db 242 ATCAAGAAGTGCTGTGAAGGCTCTTGGGGATGGCCTGTTTCGTTCCCCAG 292
RESULT 27
BE772137
LOCUS BE772137 397 bp mRNA linear EST 20-SEP-2000
DEFINITION CM4-FT0104-230600-215-a12 FT0104 Homo sapiens cDNA, mRNA sequence.
ACCESSION BE772137

VERSION BE772137.1 GI:10225795
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 397)
Dias Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Briones,M.R.,
Nagai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.F.,
Goldman,G.H., Carvalho,A.F., Matsukuma,A., Baia,G.S., Simpson,D.H.,
Brunstein,A., deOliveira,P.S., Bucher,P., Jongeneel,C.V., O'Hare
,M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and
Simpson,A.J.
TITLE Shotgun sequencing of the human transcriptome with ORF expressed
sequence tags
JOURNAL Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
MEDLINE 20202663
COMMENT Contact: Simpson A.J.G.
Laboratory of Cancer Genetics
Ludwig Institute for Cancer Research
Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
Brazil
Tel: +55-11-2704922
Fax: +55-11-2707001
Email: asimpson@ludwig.org.br
This sequence was derived from the FAPESP/LICR Human Cancer Genome
Project. This entry can be seen in the following URL
(http://www.ludwig.org.br/scripts/gethtml2.pl?t1=&t2=CM4-FT0104-230
600-215-a12&t3=2000-06-23&t4=1)
Seq primer: puc 18 forward
High quality sequence start: 4
High quality sequence stop: 397.
Features Location/Qualifiers
1..397
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_lib="FT0104"
/dev_stage="Adult"
/note="Organ: prostate tumor; Vector: puc18; Site_1: SmaI;
Site_2: SmaI; A mini-library was made by cloning products
derived from ORESTES PCR (U.S. Letters Patent application
No. 196,716 - Ludwig Institute for Cancer Research)
profiles into the pUC 18 vector. Reverse transcription of
tissue mRNA and cDNA amplification were performed under
low stringency conditions."
BASE COUNT 90 a 118 c 95 g 94 t
ORIGIN
Alignment Scores:
Pred. No.: 2.9e-29 Length: 397
Score: 327.00 Matches: 57
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 12 Gaps: 0
09-833799-13B (1-57) x BE772137 (1-397)
QY 1 AlaGlnGluProValLysGlyProValSerThrLysProGlySerCysProIleLeu 20
Db 124 GCGCAAGAGCCAGTCAAAGTCCAGTCTCCACTAAGCTGGCTCCTGCCCAATTATCTTG 183
QY 21 IleArgCysAlaMetLeuAsnProProAsnArgCysLeuLysAspThrAspCysProGly 40
Db 184 ATCCGGTGGCCCATGTTGAATCCCCCTTAACCGCTGCTTGAAGATATACTGACTGCCCAGGA 243
QY 41 IleLysLysCysCysGluGlySerCysGlyMetAlaCysPheValProGln 57
Db 244 ATCAAGAAGTGCTGTGAAGGCTCTTGGGGATGGCCTGTTTCGTTCCCCAG 294
RESULT 28
AA586974/c
LOCUS AA586974 399 bp mRNA linear EST 26-SEP-1997

DEFINITION nn69d08.s1 NCI CGAP Lar1 Homo sapiens cDNA clone IMAGE:1089135 3' similar to gb:558717_rna1 ELAFIN PRECURSOR (HUMAN); contains element THR repetitive element ;, mRNA sequence.

ACCESSION AA586974 GI:2397788

VERSION AA586974.1

KEYWORDS EST.

SOURCE human.

ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE 1 (bases 1 to 399)

AUTHORS NCI-CGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.

TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP), Tumor Gene Index

JOURNAL Unpublished (1997)

COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgaps-r@mail.nih.gov
Tissue Procurement: L. Jeffrey Medeiros, M.D., Michael R. Emmert-Buck, M.D., Ph.D.
cDNA Library Preparation: Stratagene, Inc.
cDNA Library Arrayed by: Greg Lennon, Ph.D.
DNA Sequencing by: Washington University Genome Sequencing Center
Clone distribution: NCI-CGAP clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: www-bio.llnl.gov/bbrp/image/image.html
Insert length: 523 Std Error: 0.00
Seq primer: -40m13 fwd. Et from Amersham.

FEATURES

source

1..399

Location/Qualifiers

1..399

/organism="Homo sapiens"

/db_xref="taxon:9606"

/clone="IMAGE:1089135"

/clone_lib="NCI CGAP Lar1"

/tissue_type="larynx"

/lab_host="SOLR (kanamycin resistant)"

/note="Organ: larynx; Vector: Bluescript SK-; Site 1: EcoRI; Site 2: XhoI; Cloned unidirectionally. Primer: Oligo dT. larynx. 5' adaptor sequence: 5' GAATTCGGCAG 3' 3' adaptor sequence: 5' CTCGAGTTTTTTTTTTTTTTT 3' Average insert size: 0.9 kb."

BASE COUNT 97 a 88 c 129 g 85 t

ORIGIN

Alignment Scores:

Pred. No.: 2.91e-29 Length: 399

Score: 327.00 Matches: 57

Percent Similarity: 100.00% Conservative: 0

Best Local Similarity: 100.00% Mismatches: 0

Query Match: 100.00% Indels: 0

DB: 9 Gaps: 0

09-833799-13B (1-57) x AA586974 (1-399)

QY 1 AlaGlnGluProValLysGlyProValSerThrLysProGlySerCysProIleLeu 20

Db 349 GCGCAAGAGCCAGTCAAAGGTCAGTCTCCACTAAGCCTGGCTCTGCCCCATTATCTTG 290

QY 21 IleArgCysAlaMetLeuAsnProProAsnArgCysLeuLysAspThrAspCysProGly 40

Db 289 ATCCGGTGGCCATGTTGAATCCCCCTAACCGCTGCTTGAAGATAGTACTGACTGCCAGGA 230

QY 41 IleLysLysCysCysGluGlySerCysGlyMetAlaCysPheValProGln 57

Db 229 ATCAAGAAGTGTGTGAAGGCTCTTGGCGGATGGCCTGTTTCGTTCCCCAG 179

RESULT 29

BE715592

LOCUS BE715592

DEFINITION CM4-HT0744-160600-201-b10 HT0744 Homo sapiens cDNA, mRNA sequence.

ACCESSION BE715592

VERSION BE715592.1 GI:10103857

KEYWORDS EST.

SOURCE human.

ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE 1 (bases 1 to 400)

AUTHORS Dias Neto, E., Garcia Correa, R., Verjovski-Almeida, S., Briones, M.R., Nagai, M.A., da Silva, W. Jr., Zago, M.A., Bordin, S., Costa, F.F., Goldman, G.H., Carvalho, A.F., Matsukuma, A., Baia, G.S., Simpson, D.H., Brunstein, A., deOliveira, P.S., Bucher, P., Jongeneel, C.V., O'Hare, M.J., Soares, F., Brentani, R.R., Reis, L.F., de Souza, S.J. and Simpson, A.J.

TITLE Shotgun sequencing of the human transcriptome with ORF expressed sequence tags

JOURNAL Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)

MEDLINE 20202663

COMMENT Contact: Simpson A.J.G.
Laboratory of Cancer Genetics
Ludwig Institute for Cancer Research
Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP, Brazil
Tel: +55-11-2704922
Fax: +55-11-2707001
Email: asimpson@ludwig.org.br
This sequence was derived from the FAPESP/LICR Human Cancer Genome Project. This entry can be seen in the following URL (<http://www.ludwig.org.br/scripts/gethtml2.pl?tl=ft2=CM4-HT0744-160>)
600-201-b10&t3=2000-06-16&t4=1
Seq primer: puc 18 forward
High quality sequence start: 7
High quality sequence stop: 400.

FEATURES

source

1..400

Location/Qualifiers

1..400

/organism="Homo sapiens"

/db_xref="taxon:9606"

/clone_lib="HT0744"

/dev_stage="Adult"

/note="Organ: head_neck; Vector: puc18; Site 1: SmaI; Site 2: SmaI; A mini-library was made by cloning products derived from ORESTES PCR (U.S. Letters Patent application No. 196,716 - Ludwig Institute for Cancer Research) profiles into the pUC 18 vector. Reverse transcription of tissue mRNA and cDNA amplification were performed under low stringency conditions."

BASE COUNT 90 a 115 c 100 g 95 t

ORIGIN

Alignment Scores:

Pred. No.: 2.92e-29 Length: 400

Score: 327.00 Matches: 57

Percent Similarity: 100.00% Conservative: 0

Best Local Similarity: 100.00% Mismatches: 0

Query Match: 100.00% Indels: 0

DB: 12 Gaps: 0

09-833799-13B (1-57) x BE715592 (1-400)

QY 1 AlaGlnGluProValLysGlyProValSerThrLysProGlySerCysProIleLeu 20

Db 123 GCGCAAGAGCCAGTCAAAGGTCAGTCTCCACTAAGCCTGGCTCTGCCCCATTATCTTG 182

QY 21 IleArgCysAlaMetLeuAsnProProAsnArgCysLeuLysAspThrAspCysProGly 40

Db 183 ATCCGGTGGCCATGTTGAATCCCCCTAACCGCTGCTTGAAGATAGTACTGACTGCCAGGA 242

QY 41 IleLysLysCysCysGluGlySerCysGlyMetAlaCysPheValProGln 57

Db 243 ATCAAGAAGTGTGTGAAGGCTCTTGGCGGATGGCCTGTTTCGTTCCCCAG 293

RESULT 30

BE715609/c

LOCUS BE715609

DEFINITION CM4-HT0744-160600-201-c04 HT0744 Homo sapiens cDNA, mRNA sequence.

ACCESSION BE715609

VERSION BE715609.1 GI:10103874

KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 403)
Dias Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Briones,M.R.,
Nagai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.F.,
Goldman,G.H., Carvalho,A.F., Matsukuma,A., Baia,G.S., Simpson,D.H.,
Brunstein,A., deOliveira,P.S., Bucher,P., Jongeneel,C.V., O'Hare
M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and
Simpson,A.J.
TITLE Shotgun sequencing of the human transcriptome with ORF expressed
sequence tags
JOURNAL Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
MEDLINE 20202663
COMMENT Contact: Simpson A.J.G.
Laboratory of Cancer Genetics
Ludwig Institute for Cancer Research
Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
Brazil
Tel: +55-11-2704922
Fax: +55-11-2707001
Email: asimpson@ludwig.org.br
This sequence was derived from the FAPESP/LICR Human Cancer Genome
Project. This entry can be seen in the following URL
(http://www.ludwig.org.br/scripts/gethtml2.pl?i=&t2=CM4-HT0744-160
600-201-c04&t3=2000-06-16&t4=1)
Seq primer: puc 18 forward
High quality sequence start: 12
High quality sequence stop: 403.
Location/Qualifiers
1..403
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_lib="HT0744"
/dev_stage="Adult"
/note="Organ: head neck; Vector: puc18; Site_1: SmaI;
Site_2: SmaI; A mini-library was made by cloning products
derived from ORESTES PCR (U.S. Letters patent application
No. 196,716 - Ludwig Institute for Cancer Research)
profiles into the pUC 18 vector. Reverse transcription of
tissue mRNA and cDNA amplification were performed under
low stringency conditions."

FEATURES
source

BASE COUNT 96 a 98 c 113 g 96 t
ORIGIN
Alignment Scores:
Pred. No.: 2.95e-29 Length: 403
Score: 327.00 Matches: 57
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 12 Gaps: 0
09-833799-13B (1-57) x BE715609 (1-403)

QY 1 AlaGlnGluProValLysGlyProValSerThrLysProGlySerCysProIleIleLeu 20
Db 272 GCGAAGAGCCAGTCAAAGGTCAGTCTCCACTAAGCCTGGCTCCTGCCCATTTATCTTG 213
QY 21 IleArgCysAlaMetLeuAsnProProAsnArgCysLeuLysAspThrAspCysProGly 40
Db 212 ATCCGGTGGCCCATGTTGAATCCCCCTAACCGCTGCTTGAAAGATACTGACTGCCCAGGA 153
QY 41 IleLysLysCysCysGluGlySerCysGlyMetAlaCysPheValProGln 57
Db 152 ATCAAGAAGTGTGTGAAGGCTCTTGGCGGATGGCCTGTTTCGTTCCCCAG 102

RESULT 31
AI874186/c
LOCUS AI874186 405 bp mRNA linear EST 07-MAR-2000
DEFINITION wm50a02.x1 NCI_CGAP_Ut4 Homo sapiens cDNA clone IMAGE:2439338 3'

ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT

similar to gb:S58717_rnal ELAFIN PRECURSOR (HUMAN);, mRNA sequence.
AI874186
AI874186.1 GI:5548235
EST.
human.
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 405)
NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
Unpublished (1997)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
Tissue Procurement: Christopher Moskaluk, M.D., Ph.D., Michael R.
Emmert-Buck, M.D., Ph.D.
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: Greg Lennon, Ph.D.
DNA Sequencing by: Washington University Genome Sequencing Center
Clone distribution: NCI-CGAP clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
www-bio.llnl.gov/bbrp/image/image.html
Insert Length: 544 Std Error: 0.00
Seq primer: -40UP from Gibco.
Location/Qualifiers
1..405
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_lib="IMAGE:2439338"
/clone_lib="NCI_CGAP_Ut4"
/tissue_type="serous papillary carcinoma, high grade, 2
pooled tumors"
/lab_host="DH10B"
/note="Organ: uterus; Vector: pCMV-SPORT6; Site_1: SmaI;
Site_2: NotI; Cloned unidirectionally. Primer: Oligo dt.
Average insert size 1.48 kb. Life Technologies catalog #:
11542-016"

FEATURES
source

BASE COUNT 96 a 84 c 125 g 100 t
ORIGIN
Alignment Scores:
Pred. No.: 2.96e-29 Length: 405
Score: 327.00 Matches: 57
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 9 Gaps: 0
09-833799-13B (1-57) x AI874186 (1-405)

QY 1 AlaGlnGluProValLysGlyProValSerThrLysProGlySerCysProIleIleLeu 20
Db 375 GCGAAGAGCCAGTCAAAGGTCAGTCTCCACTAAGCCTGGCTCCTGCCCATTTATCTTG 316
QY 21 IleArgCysAlaMetLeuAsnProProAsnArgCysLeuLysAspThrAspCysProGly 40
Db 315 ATCCGGTGGCCCATGTTGAATCCCCCTAACCGCTGCTTGAAAGATACTGACTGCCCAGGA 256
QY 41 IleLysLysCysCysGluGlySerCysGlyMetAlaCysPheValProGln 57
Db 255 ATCAAGAAGTGTGTGAAGGCTCTTGGCGGATGGCCTGTTTCGTTCCCCAG 205

RESULT 32
AW137392/c
LOCUS AW137392 406 bp mRNA linear EST 29-OCT-1999
DEFINITION UI-H-B11-acs-b-12-0-UI.s1 NCI_CGAP_Sub3 Homo sapiens cDNA clone
IMAGE:2715238 3', mRNA sequence.
ACCESSION AW137392
VERSION AW137392.1 GI:6141710
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens

09-833799-13B (1-57) x BE715604 (1-406)

Qy 1 AlaGlnGluProValLysGlyProValSerThrLysProGlySerCysProIleIleLeu 20
Db 133 GCGCAAGAGCCAGTCAAAGGTCAGTCTCCACTAAGCCTGGCTCTGCCCATTAATCTG 192

BASE COUNT	103 a	92 c	134 g	88 t	1 others
ORIGIN					
Alignment Scores:					
Pred. No.:	3.08e-29				418
Score:	327.00				57
Percent Similarity:	100.00%				0
Best Local Similarity:	100.00%				0
Query Match:	100.00%				0
DB:	9				0

09-833799-13B (1-57) x AI392753 (1-418)

Qy 1 AlaGlnGluProValLysGlyProValSerThrLysProGlySerCysProIleLeu 20
|||
Db 349 GCGCAAGAGCCAGTCAAAGGTCACAGTCTCCACTAAGCCTGGTCTCTGCCCATATATCTG 290

BASE COUNT	103 a	92 c	135 g	89 t
ORIGIN				
Alignment Scores:				
Pred. NO.:	3.09e-29	Length:	419	
Score:	327.00	Matches:	57	
Percent Similarity:	100.00%	Conservative:	0	
Best Local Similarity:	100.00%	Mismatches:	0	
Query Match:	100.00%	Indels:	0	
DB:	9	Gaps:	0	

09-833799-13B (1-57) x AA586943 (1-419)

Qy 1 AlaGlnGluProValLysGlyProValSerThrLysProGlySerCysProIleIleLeu 20

```
Db 349 GCGCAAGAGCCAGTCAAAGGTCCAGTCTCCACTAAGCCTGGCTCCTGCCCATATCTTG 290
QY 21 IleArgCysAlaMetLeuAsnProProAsnArgCysLeuLysAspThrAspCysProGly 40
Db 289 ATCCGGTGCAGCCATGTTGAATCCCCCTAACCCGCTGCTTGAAGATACTGACTGCCAGGA 230

QY 41 IleLysLysCysCysGluGlySerCysGlyMetAlaCysPheValProGln 57
Db 229 ATCAAGAAGTGTGTGAAGGCTCTTGGGGAAGGCTGTTTCGTTCCCGAG 179

RESULT 36
AW082097/c
LOCUS
DEFINITION
xb60d08.x1 NCI CGAP Eso2 Homo sapiens cDNA clone IMAGE:2580687 3'
similar to gb:S58717_xnal ELAFIN PRECURSOR (HUMAN); contains element
MER28 MER28 repetitive element ;, mRNA sequence.
AW082097
ACCESSION
AW082097.1 GI:6037249
VERSION
EST.
KEYWORDS
SOURCE
human.
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
1 (bases 1 to 422)
AUTHORS
NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
TITLE
National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
JOURNAL
Unpublished (1997)
COMMENT
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-re@mail.nih.gov
Tissue Procurement: Nan Hu, M.D., Ph.D., Mark Roth, M.D., Phillip
Taylor, M.D., Michael R. Emmert-Buck, M.D., Ph.D.
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: Greg Lennon, Ph.D.
DNA Sequencing by: Washington University Genome Sequencing Center
Clone distribution: NCI-CGAP clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
www-bio.llnl.gov/bbrp/image/image.html

Possible reversed clone: polyT not found
Seq primer: -40UP from Gibco
High quality sequence stop: 407.
FEATURES
Location/Qualifiers
1..422
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:2580687"
/clone_lib="NCI CGAP Eso2"
/tissue_type="squamous cell carcinoma"
/lab_host="DH10B"
/note="Organ: esophagus; Vector: pCMV-sport6; Site 1: SalI
; Site 2: NotI; Cloned unidirectionally. Primer: Oligo
dT. Average insert size 1.1 kb. Life Technologies catalog
#: 11502-010"
BASE COUNT 98 a 88 c 129 g 107 t
ORIGIN

Alignment Scores:
Pred. No.: 3.11e-29 Length: 422
Score: 327.00 Matches: 57
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 10 Gaps: 0

09-833799-13B (1-57) x AW082097 (1-422)

QY 1 AlaGlnGluProValLysGlyProValSerThrLysProGlySerCysProIleLeu 20
Db 371 GCGCAAGAGCCAGTCAAAGGTCCAGTCTCCACTAAGCCTGGCTCCTGCCCATATCTTG 312
QY 21 IleArgCysAlaMetLeuAsnProProAsnArgCysLeuLysAspThrAspCysProGly 40
```

```
Db 311 ATCCGGTGCAGCCATGTTGAATCCCCCTAACCGCTGCTTGAAGATACTGACTGCCAGGA 252
QY 41 IleLysLysCysCysGluGlySerCysGlyMetAlaCysPheValProGln 57
Db 251 ATCAAGAAGTGTGTGAAGGCTCTTGGGGAAGGCTGTTTCGTTCCCGAG 201

RESULT 37
BE772154/c
LOCUS
DEFINITION
CM4-FT0104-230600-215-b04 FT0104 Homo sapiens cDNA, mRNA sequence.
ACCESSION
BE772154
VERSION
BE772154.1 GI:10225812
KEYWORDS
EST.
SOURCE
human.
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
1 (bases 1 to 428)
AUTHORS
Dias Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Briones,M.R.,
Nagai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.F.,
Goldman,G.H., Carvalho,A.F., Matsukuma,A., Baia,G.S., Simpson,D.H.,
Brunstein,A., deOliveira,P.S., Bucher,P., Jongeneel,C.V., O'Hare
,M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and
Simpson,A.J.
TITLE
Shotgun sequencing of the human transcriptome with ORF expressed
sequence tags
JOURNAL
Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
MEDLINE
20202663
COMMENT
Contact: Simpson A.J.G.
Laboratory of Cancer Genetics
Ludwig Institute for Cancer Research
Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
Brazil
Tel: +55-11-2704922
Fax: +55-11-2707001
Email: asimpson@ludwig.org.br
This sequence was derived from the FAPESP/LICR Human Cancer Genome
Project. This entry can be seen in the following URL
(http://www.ludwig.org.br/scripts/gethtml2.pl?tl=&t2=CM4-FT0104-230
600-215-b04&t3=2000-06-23&t4=1)
Seq primer: puc 18 forward
High quality sequence start: 13
High quality sequence stop: 428.
FEATURES
Location/Qualifiers
1..428
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_lib="FT0104"
/dev_stage="Adult"
/note="Organ: prostate tumor; Vector: puc18; Site 1: SmaI;
Site 2: SmaI; A mini-library was made by cloning products
derived from ORESTES PCR (U.S. Letters Patent application
No. 196,716 - Ludwig Institute for Cancer Research)
profiles into the pUC 18 vector. Reverse transcription of
tissue mRNA and cDNA amplification were performed under
low stringency conditions."
BASE COUNT 101 a 99 c 131 g 97 t
ORIGIN

Alignment Scores:
Pred. No.: 3.16e-29 Length: 428
Score: 327.00 Matches: 57
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 12 Gaps: 0

09-833799-13B (1-57) x BE772154 (1-428)

QY 1 AlaGlnGluProValLysGlyProValSerThrLysProGlySerCysProIleLeu 20
Db 333 GCGCAAGAGCCAGTCAAAGGTCCAGTCTCCACTAAGCCTGGCTCCTGCCCATATCTTG 274
```

QY 21 IleArgCysAlaMetLeuAsnProProAsnArgCysLeuLysAspThrAspCysProGly 40
|||||
Db 273 ATCCGGTGGCCCATGTTGAATCCCCCTAACCGCTGCTTGAAGATAGTACTGACTGCCCAGGA 214

QY 41 IleLysLysCysCysGluGlySerCysGlyMetAlaCysPheValProGln 57
|||||
Db 213 ATCAAGAAAGTGCTGTGAAGGCTCTTGGCGGATGGCCTGTTTCGTTCCCCAG 163

RESULT 38
AI283910/C
LOCUS
DEFINITION Qt66a10.x1 NCI_CGAP_Eso2 Homo sapiens cDNA clone IMAGE:1960218 3'
similar to gb:S58717_rnal ELAFIN PRECURSOR (HUMAN); contains element
THR repetitive element ;, mRNA sequence.

ACCESSION AI283910
VERSION AI283910.1 GI:3922143
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 429)
NCI-CGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.
National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
Unpublished (1997)
Contact: Robert Strausberg, Ph.D.
Email: cgaps-r@mail.nih.gov
Tissue Procurement: Nan Hu, M.D., Ph.D., Mark Roth, M.D., Phillip
Taylor, M.D., Michael R. Emmert-Buck, M.D., Ph.D.
CDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: Greg Lennon, Ph.D.
DNA Sequencing by: Washington University Genome Sequencing Center
Clone distribution: NCI-CGAP clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
www-bio.llnl.gov/bbrp/image/image.html
Insert Length: 626 Std Error: 0.00
Seq primer: -40UP from Gibco
High quality sequence stop: 428.
Location/Qualifiers
1. .429
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:1960218"
/clone_lib="NCI_CGAP_Eso2"
/tissue_type="squamous cell carcinoma"
/lab_host="DH10B"
/note="Organ: esophagus; Vector: pCMV-SPORT6; Site 1: Sali
; Site 2: NotI; Cloned unidirectionally. Primer: Oligo
dT. Average insert size 1.1 kb. Life Technologies catalog
#: 11502-010"

BASE COUNT 102 a 95 c 136 g 93 t 3 others
ORIGIN

Alignment Scores:
Pred. No.: 3.17e-29 Length: 429
Score: 327.00 Matches: 57
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 9 Gaps: 0

09-833799-13B (1-57) x AI283910 (1-429)

QY 1 AlaGlnGluProValLysGlyProValSerThrLysProGlySerCysProIleLeu 20
|||||
Db 350 GCGCAAGAGCCAGTCAAGAGTCCAGTCTCCACTAAGCCTGGCTCCTGCCCATATCTTG 291

QY 21 IleArgCysAlaMetLeuAsnProProAsnArgCysLeuLysAspThrAspCysProGly 40
|||||
Db 290 ATCCGGTGGCCCATGTTGAATCCCCCTAACCGCTGCTTGAAGATAGTACTGACTGCCCAGGA 231

QY 41 IleLysLysCysCysGluGlySerCysGlyMetAlaCysPheValProGln 57
|||||
Db 230 ATCAAGAAAGTGCTGTGAAGGCTCTTGGCGGATGGCCTGTTTCGTTCCCCAG 180

RESULT 39
BF836224
LOCUS
DEFINITION QV3-HT1016-171100-474-h01 HT1016 Homo sapiens cDNA, mRNA sequence.
ACCESSION BF836224
VERSION BF836224.1 GI:12187935
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 437)
Dias Neto, E., Garcia Correa, R., Verjovski-Almeida, S., Briones, M.R.,
Nagai, M.A., da Silva, W. Jr., Zago, M.A., Bordin, S., Costa, F.F.,
Goldman, G.H., Carvalho, A.F., Matsukuma, A., Baia, G.S., Simpson, D.H.,
Brunstein, A., deOliveira, P.S., Bucher, P., Jongeneel, C.V., O'Hare,
M.J., Soares, F., Brentani, R.R., Reis, L.F., de Souza, S.J. and
Simpson, A.J.
Shotgun sequencing of the human transcriptome with ORF expressed
sequence tags
Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
20202663
Contact: Simpson A.J.G.
Laboratory of Cancer Genetics
Ludwig Institute for Cancer Research
Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
Brazil
Tel: +55-11-2704922
Fax: +55-11-2707001
Email: asimpson@ludwig.org.br
This sequence was derived from the FAPESP/LICR Human Cancer Genome
Project. This entry can be seen in the following URL
(<http://www.ludwig.org.br/scripts/gethtml2.pl?ti=QV3&t2=QV3-HT1016-171100-474-h01&t3=2000-11-17&t4=1>)
Seq primer: puc 18 forward
High quality sequence start: 13
High quality sequence stop: 437.
Location/Qualifiers
1. .437
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_lib="HT1016"
/dev_stage="Adult"
/note="Organ: head neck; Vector: puc18; Site 1: SmaI;
Site 2: SmaI; A mini-library was made by cloning products
derived from ORESTES PCR (U.S. Letters Patent application
No. 196,716 - Ludwig Institute for Cancer Research)
profiles into the pUC 18 vector. Reverse transcription of
tissue mRNA and cDNA amplification were performed under
low stringency conditions."

BASE COUNT 99 a 120 c 118 g 100 t
ORIGIN

Alignment Scores:
Pred. No.: 3.24e-29 Length: 437
Score: 327.00 Matches: 57
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 12 Gaps: 0

09-833799-13B (1-57) x BF836224 (1-437)

QY 1 AlaGlnGluProValLysGlyProValSerThrLysProGlySerCysProIleLeu 20
|||||
Db 198 GCGCAAGAGCCAGTCAAGAGTCCAGTCTCCACTAAGCCTGGCTCCTGCCCATATCTTG 257

QY 21 IleArgCysAlaMetLeuAsnProProAsnArgCysLeuLysAspThrAspCysProGly 40
|||||

```
Db 258 ATCCGGTGGCCATGTTGAATCCCTTAACCGCTGCTTGAAAGATACTGACTGCCAGGA 317
QY 41 IleLysLysCysCysGluGlySerCysGlyMetAlaCysPheValProGln 57
Db 318 ATCAAGAAGTGTGTGAAGCTCTTGCGGGATGGCTGTTTCGTTCCCCAG 368

RESULT 40
AW001880/c
LOCUS
DEFINITION
IMAGE:2513867 3' similar to gb:S58717 rnal ELAFIN PRECURSOR (HUMAN
);contains element MER28 repetitive element ;, mRNA sequence.
ACCESSION
AW001880
VERSION
AW001880.1 GI:5848796
KEYWORDS
EST.
SOURCE
human.
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 438)
Dias Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Briones,M.R.,
Nagai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.F.,
Goldman,G.H., Carvalho,A.F., Matsukuma,A., Baia,G.S., Simpson,D.H.,
Brunstein,A., deOliveira,P.S., Bucher,P., Jongeneel,C.V., O'Hare
,M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and
Simpson,A.J.
Shotgun sequencing of the human transcriptome with ORF expressed
sequence tags
Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
20202663
Contact: Simpson A.J.G.
Laboratory of Cancer Genetics
Ludwig Institute for Cancer Research
Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
Brazil
Tel: +55-11-2704922
Fax: +55-11-2707001
Email: asimpson@ludwig.org.br
This sequence was derived from the FAPESP/LICR Human Cancer Genome
Project. This entry can be seen in the following URL
(http://www.ludwig.org.br/scripts/gethtml2.pl?tl=et2=CM4-HT0744-160
600-201-a04&t3=2000-06-16&t4=1)
Seq primer: puc 18 forward
High quality sequence start: 24
High quality sequence stop: 437.
FEATURES
source
1..438
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_lib="IMAGE:2513867"
/clone_lib="Soares_thymus_NHFTh"
/dev_stage="fetal"
/lab_host="DH10B (phage-resistant)"
/note="Organ: thymus, pooled; Vector: pT7T3D-Pac
(Pharmacia) with a modified polylinker; Site 1: Not I;
Site 2: Eco RI; 1st strand cDNA was primed with a Not I -
oligo(dT) primer [5,
TGTTACCAATCTGAAGTGGGAGCGCGCAACGTTTTTTTTTTTTTTT 3'],
double-stranded cDNA was ligated to Eco RI adaptors
(Pharmacia), digested with Not I and cloned into the Not I
and Eco RI sites of the modified pT7T3 vector. Library
went through one round of normalization. Library
constructed by Bento Soares and M. Fatima Bonaldo. "
BASE COUNT 106 a 92 c 125 g 115 t
ORIGIN
Alignment Scores:
Pred. No.: 3.25e-29 Length: 438
Score: 327.00 Matches: 57
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
Gaps: 10
DB:
09-833799-13B (1-57) x AW001880 (1-438)
QY 1 AlaGlnGluProValLysGlyProValSerThrLysProGlySerCysProIleIleLeu 20
Db 375 GCGCAAGAGCAGTCAAAGGTCACACTAAGCCTGGCTCCTGCCCATTAATTG 316
QY 21 IleArgCysAlaMetLeuAsnProProAsnArgCysLeuLysAspThrAspCysProGly 40
Db 315 ATCCGGTGGCCATGTTGAATCCCCCTAACCGCTGCTTGAAAGATACTGACTGCCAGGA 256
QY 41 IleLysLysCysCysGluGlySerCysGlyMetAlaCysPheValProGln 57
Db 255 ATCAAAAAGTGTGTGAAGGTTCTTGCGGGATGGCTGTTTCGTTCCCCAG 205
```

```
RESULT 41
BE715601/c
LOCUS
DEFINITION
CM4-HT0744-160600-201-a04 HT0744 Homo sapiens cDNA, mRNA sequence.
ACCESSION
BE715601
VERSION
BE715601.1 GI:10103866
KEYWORDS
EST.
SOURCE
human.
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 438)
Dias Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Briones,M.R.,
Nagai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.F.,
Goldman,G.H., Carvalho,A.F., Matsukuma,A., Baia,G.S., Simpson,D.H.,
Brunstein,A., deOliveira,P.S., Bucher,P., Jongeneel,C.V., O'Hare
,M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and
Simpson,A.J.
Shotgun sequencing of the human transcriptome with ORF expressed
sequence tags
Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
20202663
Contact: Simpson A.J.G.
Laboratory of Cancer Genetics
Ludwig Institute for Cancer Research
Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
Brazil
Tel: +55-11-2704922
Fax: +55-11-2707001
Email: asimpson@ludwig.org.br
This sequence was derived from the FAPESP/LICR Human Cancer Genome
Project. This entry can be seen in the following URL
(http://www.ludwig.org.br/scripts/gethtml2.pl?tl=et2=CM4-HT0744-160
600-201-a04&t3=2000-06-16&t4=1)
Seq primer: puc 18 forward
High quality sequence start: 24
High quality sequence stop: 437.
FEATURES
source
1..438
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_lib="HT0744"
/dev_stage="Adult"
/note="Organ: head_neck; Vector: puc18; Site 1: SmaI;
Site 2: SmaI; A mini-library was made by cloning products
derived from ORESTES PCR (U.S. Letters Patent application
No. 196,716 - Ludwig Institute for Cancer Research)
profiles into the pUC 18 vector. Reverse transcription of
tissue mRNA and cDNA amplification were performed under
low stringency conditions."
BASE COUNT 104 a 102 c 132 g 100 t
ORIGIN
Alignment Scores:
Pred. No.: 3.25e-29 Length: 438
Score: 327.00 Matches: 57
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
Gaps: 12
DB:
09-833799-13B (1-57) x BE715601 (1-438)
QY 1 AlaGlnGluProValLysGlyProValSerThrLysProGlySerCysProIleIleLeu 20
Db 327 GCGCAAGAGCAGTCAAAGGTCACACTAAGCCTGGCTCCTGCCCATTAATTG 268
QY 21 IleArgCysAlaMetLeuAsnProProAsnArgCysLeuLysAspThrAspCysProGly 40
Db 267 ATCCGGTGGCCATGTTGAATCCCCCTAACCGCTGCTTGAAAGATACTGACTGCCAGGA 208
QY 41 IleLysLysCysCysGluGlySerCysGlyMetAlaCysPheValProGln 57
```

Db 207 ATCAAGAAGTGCTGTGAAGGCTCTTGGCGGATGGCCTGTTTCGTTCCCCAG 157
RESULT 42
BF836743
LOCUS BF836743 441 bp mRNA linear EST 13-JAN-2001
DEFINITION CM2-HT0969-181100-509-b06 HT0969 Homo sapiens cDNA, mRNA sequence.
ACCESSION BF836743
VERSION BF836743.1 GI:12188790
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 441)
AUTHORS Dias Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Briones,M.R.,
Nagai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.F.,
Goldman,G.H., Carvalho,A.F., Matsukuma,A., Baia,G.S., Simpson,D.H.,
Brunstein,A., deOliveira,P.S., Bucher,P., Jongeneel,C.V., O'Hare
,M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and
Simpson,A.J.
TITLE Shotgun sequencing of the human transcriptome with ORF expressed
sequence tags
JOURNAL Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
MEDLINE 20202663
COMMENT Contact: Simpson A.J.G.
Laboratory of Cancer Genetics
Ludwig Institute for Cancer Research
Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
Brazil
Tel: +55-11-2704922
Fax: +55-11-2707001
Email: asimpson@ludwig.org.br
This sequence was derived from the FAPESP/LICR Human Cancer Genome
Project. This entry can be seen in the following URL
(http://www.ludwig.org.br/scripts/gethtml2.pl?tl=CM2&t2=CM2-HT0969-
181100-509-b06&t3=2000-11-18&t4=1)
Seq primer: puc 18 forward
High quality sequence start: 8
High quality sequence stop: 441.
FEATURES
source
1..441
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_lib="HT0969"
/dev_stage="Adult"
/note="Organ: head neck; Vector: puc18; Site_1: SmaI;
Site_2: SmaI; A mini-library was made by cloning products
derived from ORESTES PCR (U.S. Letters Patent application
No. 196,716 - Ludwig Institute for Cancer Research)
profiles into the pUC 18 vector. Reverse transcription of
tissue mRNA and cDNA amplification were performed under
low stringency conditions."
BASE COUNT 94 a 131 c 109 g 107 t
ORIGIN
Alignment Scores:
Pred. No.: 3.28e-29 Length: 441
Score: 327.00 Matches: 57
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 12 Gaps: 0
09-833799-13B (1-57) x BF836743 (1-441)
QY 1 AlaGlnGluProValLysGlyProValSerThrLysProGlySerCysProIleLeu 20
|||||
Db 148 GCGCAAGAGCCAGTCAAAAGGTCAGTCTCCACTAAGCCTGGCTCCTGCCCATATATCTTG 207
QY 21 IleArgCysAlaMetLeuAsnProProAsnArgCysLeuLysAspThrAspCysProGly 40
|||||
Db 208 ATCCGGTGGCCCATGTTGAATCCCCCTAACCGCTGCTGAAAGATACTGACTGCCAGGA 267

QY 41 IleLysLysCysCysGluGlySerCysGlyMetAlaCysPheValProGln 57
|||||
Db 268 ATCAAGAAGTGCTGTGAAGGCTCTTGGCGGATGGCCTGTTTCGTTCCCCAG 318
RESULT 43
AI582329/c
LOCUS AI582329 448 bp mRNA linear EST 14-DEC-1999
DEFINITION tq66g09.x1 NCI CGAP Lu19 Homo sapiens cDNA clone IMAGE:2213824 3'
similar to gb:S58717_rnal ELAFIN PRECURSOR (HUMAN);contains element
THR repetitive element ;, mRNA sequence.
ACCESSION AI582329
VERSION AI582329.1 GI:4568226
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 448)
AUTHORS NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
JOURNAL Unpublished (1997)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: Christopher Moskaluk, M.D., Ph.D., Michael R.
Emmert-Buck, M.D., Ph.D.
CDNA Library Preparation: M. Bento Soares, Ph.D.
CDNA Library Arrayed by: Greg Lennon, Ph.D.
DNA Sequencing by: Washington University Genome Sequencing Center
Clone distribution: NCI-CGAP clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
www-bio.llnl.gov/bbrp/image/image.html
Insert Length: 598 Std Error: 0.00
Seq primer: -40UP from Gibco
POLYA=No.
FEATURES
source
1..448
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:2213824"
/clone_lib="NCI-CGAP Lu19"
/tissue_type="squamous cell carcinoma, poorly
differentiated (4 pooled tumors, including primary and
metastatic)"
/dev_stage="adult"
/lab_host="DH10B (phage-resistant)"
/note="Organ: lung; Vector: pT7T3D-Pac (Pharmacia) with a
modified polylinker; 1st strand cDNA was prepared from
pooled lung tumor tissue, and was then primed with a Not I
- oligo(dT) primer. Double-stranded cDNA was ligated to
Eco RI adaptors (Pharmacia), digested with Not I and
cloned into the Not I and Eco RI sites of the modified
pT7T3 vector. Library went through one round of
normalization. Library constructed by Bento Soares and M.
Fatima Bonaldo."
BASE COUNT 105 a 96 c 138 g 109 t
ORIGIN
Alignment Scores:
Pred. No.: 3.34e-29 Length: 448
Score: 327.00 Matches: 57
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 9 Gaps: 0
09-833799-13B (1-57) x AI582329 (1-448)
QY 1 AlaGlnGluProValLysGlyProValSerThrLysProGlySerCysProIleLeu 20
|||||
Db 364 GCGCAAGAGCCAGTCAAAAGGTCAGTCTCCACTAAGCCTGGCTCCTGCCCATATATCTTG 305
QY 21 IleArgCysAlaMetLeuAsnProProAsnArgCysLeuLysAspThrAspCysProGly 40

```
Db 304 ATCCGGTGGCCATGTTGAATCCCCCTAACCGCTGCTTGAAGATACTGACTGCCAGGA 245
|||||
QY 41 IleLysLysCysCysGluGlySerCysGlyMetAlaCysPheValProGln 57
|||||
Db 244 ATCAAGAGTGTGTGAAGGCTCTTGGCGGATGGCCTGTTTCGTTCCCCAG 194
|||||

RESULT 44
AI459240/c
LOCUS AI459240 453 bp mRNA linear EST 14-APR-1999
DEFINITION tk11c09.x1 NCI CGAP Lu24 Homo sapiens cDNA clone IMAGE:2150704 3'
similar to gb:S58717_rnal ELAFIN PRECURSOR (HUMAN);contains element
THR repetitive element ;, mRNA sequence.
ACCESSION AI459240
VERSION AI459240.1 GI:4311819
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 453)
AUTHORS Dias Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Briones,M.R.,
Nagai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.F.,
Goldman,G.H., Carvalho,A.F., Matsukuma,A., Baia,G.S., Simpson,D.H.,
Brunstein,A., deOliveira,P.S., Bucher,P., Jongeneel,C.V., O'Hare
,M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and
Simpson,A.J.
TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
JOURNAL Unpublished (1997)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: Christopher Moskaluk, M.D., Ph.D., Michael R.
Emmert-Buck, M.D., Ph.D.
cDNA Library Preparation: M. Bento Soares, Ph.D.
DNA Library Arrayed by: Greg Lennon, Ph.D.
DNA Sequencing by: Washington University Genome Sequencing Center
Clone distribution: NCI-CGAP clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
www-bio.llnl.gov/bbrp/image/image.html
Insert Length: 554 Std Error: 0.00
Seq primer: -40UP from Gibco.
Location/Qualifiers
1. 453
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:2150704"
/clone_lib="NCI CGAP Lu24"
/tissue_type="carcinoid"
/lab_host="DH10B"
/note="Organ: lung; Vector: pT7T3D-Pac (Pharmacia) with a
modified polylinker; Plasmid DNA from the normalized
library NCI CGAP Lu5 was prepared, and ss circles were
made in vitro. Following HAP purification, this DNA was
used as tracer in a subtractive hybridization reaction.
The driver was PCR-amplified cDNAs from a pool of 5,000
clones made from the same library (cloneIDs
1414920-1417991 and 1520904-1522439). Subtraction by Bento
Soares and M. Fatima Bonaldo."
BASE COUNT 107 a 96 c 137 g 113 t
ORIGIN

Alignment Scores:
Pred. No.: 3.38e-29 Length: 453
Score: 327.00 Matches: 57
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 9 Gaps: 0

09-833799-13B (1-57) x AI459240 (1-453)

QY 1 AlaGlnGluProValLysGlyProValSerThrLysProGlySerCysProIleIleLeu 20
|||||
Db 366 GCGCAAGAGCCAGTCAAAGGTCACACTAAGCCTGGCTCCTGCCCATTTATCTTG 307
|||||
QY 21 IleArgCysAlaMetLeuAsnProProAsnArgCysLeuLysAspThrAspCysProGly 40
|||||
```

```
Db 306 ATCCGGTGGCCATGTTGAATCCCCCTAACCGCTGCTTGAAGATACTGACTGCCAGGA 247
|||||
QY 41 IleLysLysCysCysGluGlySerCysGlyMetAlaCysPheValProGln 57
|||||
Db 246 ATCAAGAGTGTGTGAAGGCTCTTGGCGGATGGCCTGTTTCGTTCCCCAG 196
|||||

RESULT 45
BF824934
LOCUS BF824934 465 bp mRNA linear EST 13-JAN-2001
DEFINITION IL0-HN0039-101100-500-b06 HN0039 Homo sapiens cDNA, mRNA sequence.
ACCESSION BF824934
VERSION BF824934.1 GI:12167001
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 465)
AUTHORS Dias Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Briones,M.R.,
Nagai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.F.,
Goldman,G.H., Carvalho,A.F., Matsukuma,A., Baia,G.S., Simpson,D.H.,
Brunstein,A., deOliveira,P.S., Bucher,P., Jongeneel,C.V., O'Hare
,M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and
Simpson,A.J.
TITLE Shotgun sequencing of the human transcriptome with ORF expressed
sequence tags
JOURNAL Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
MEDLINE 20202663
COMMENT Contact: Simpson A.J.G.
Laboratory of Cancer Genetics
Ludwig Institute for Cancer Research
Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
Brazil
Tel: +55-11-2704922
Fax: +55-11-2707001
Email: asimpson@ludwig.org.br
This sequence was derived from the FAPESP/LICR Human Cancer Genome
Project. This entry can be seen in the following URL
(http://www.ludwig.org.br/scripts/gethtml2.pl?t1=IL0&t2=IL0-HN0039-
101100-500-b06&t3=2000-11-10&t4=1)
Seq primer: puc 18 forward
High quality sequence stop: 465.
Location/Qualifiers
1. 465
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_lib="HN0039"
/dev_stage="Adult"
/note="Organ: head normal; Vector: puc18; Site 1: Smal;
Site 2: Smal; A mini-library was made by cloning products
derived from ORESTES PCR (U.S. Letters Patent application
No. 196,716 - Ludwig Institute for Cancer Research)
profiles into the pUC 18 vector. Reverse transcription of
tissue mRNA and cDNA amplification were performed under
low stringency conditions."
BASE COUNT 98 a 137 c 117 g 113 t
ORIGIN

Alignment Scores:
Pred. No.: 3.49e-29 Length: 465
Score: 327.00 Matches: 57
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 12 Gaps: 0

09-833799-13B (1-57) x BF824934 (1-465)

QY 1 AlaGlnGluProValLysGlyProValSerThrLysProGlySerCysProIleIleLeu 20
|||||
Db 183 GCGCAAGAGCCAGTCAAAGGTCACACTAAGCCTGGCTCCTGCCCATTTATCTTG 242
|||||
QY 21 IleArgCysAlaMetLeuAsnProProAsnArgCysLeuLysAspThrAspCysProGly 40
|||||
```

Db 243 ATCCGGTGGCCCATGTTGAATCCCCTAACCGTGCTTGAAAGATACTGACTGCCAGGA 302
QY 41 IleLysLysCysCysGluGlySerCysGlyMetAlaCysPheValProGln 57
Db 303 ATCAAGAAGTGCTGTGAAGGCTCTGCGGGATGGCCTGTTTCGTTCCCCAG 353
RESULT 46
BF824930
LOCUS BF824930 466 bp mRNA linear EST 13-JAN-2001
DEFINITION ILO-HN0039-101100-500-a05 HN0039 Homo sapiens cDNA, mRNA sequence.
ACCESSION BF824930
VERSION BF824930.1 GI:12166995
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 466)
AUTHORS Nagai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.F., Goldman,G.H., Carvalho,A.F., Matsukuma,A., Baia,G.S., Simpson,D.H., Brunstein,A., deOliveira,P.S., Bucher,P., Jongeneel,C.V., O'Hare ,M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and Simpson,A.J.
TITLE Shotgun sequencing of the human transcriptome with ORF expressed sequence tags
JOURNAL Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
MEDLINE 20202663
COMMENT Contact: Simpson A.J.G.
Laboratory of Cancer Genetics
Ludwig Institute for Cancer Research
Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP, Brazil
Tel: +55-11-2704922
Fax: +55-11-2707001
Email: asimpson@ludwig.org.br
This sequence was derived from the FAPESP/LICR Human Cancer Genome Project. This entry can be seen in the following URL (http://www.ludwig.org.br/scripts/gethtml2.pl?tl=ILO&t2=ILO-HN0039-101100-500-a05&t3=2000-11-10&t4=1)
Seq primer: puc 18 forward
High quality sequence stop: 466.
Location/Qualifiers
1. .466
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_lib="HN0039"
/dev_stage="Adult"
/note="Organ: head normal; Vector: puc18; Site_1: SmaI; Site_2: SmaI; A mini-library was made by cloning products derived from ORESTES PCR (U.S. Letters Patent application No. 196,716 - Ludwig Institute for Cancer Research) profiles into the pUC 18 vector. Reverse transcription of tissue mRNA and cDNA amplification were performed under low stringency conditions."
BASE COUNT 98 a 138 c 117 g 113 t
ORIGIN
Alignment Scores:
Pred. No.: 3.5e-29 Length: 466
Score: 327.00 Matches: 57
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 12 Gaps: 0
09-833799-13B (1-57) x BF824930 (1-466)
QY 1 AlaGlnGluProValSerThrLysProGlySerCysProIleLeu 20
Db 184 GCGAAGAGCCGGTCAAGGTCAGTCTCCACTAAGCCTGGCTCCTGCCCATATATCTTG 243

QY 21 IleArgCysAlaMetLeuAsnProProAsnArgCysLeuLysAspThrAspCysProGly 40
Db 244 ATCCGGTGGCCCATGTTGAATCCCCTAACCGTGCTTGAAAGATACTGACTGCCAGGA 303
QY 41 IleLysLysCysCysGluGlySerCysGlyMetAlaCysPheValProGln 57
Db 304 ATCAAGAAGTGCTGTGAAGGCTCTGCGGGATGGCCTGTTTCGTTCCCCAG 354
RESULT 47
BF824935
LOCUS BF824935 466 bp mRNA linear EST 13-JAN-2001
DEFINITION ILO-HN0039-101100-500-c03 HN0039 Homo sapiens cDNA, mRNA sequence.
ACCESSION BF824935
VERSION BF824935.1 GI:12167002
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 466)
AUTHORS Dias Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Briones,M.R., Nagai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.F., Goldman,G.H., Carvalho,A.F., Matsukuma,A., Baia,G.S., Simpson,D.H., Brunstein,A., deOliveira,P.S., Bucher,P., Jongeneel,C.V., O'Hare ,M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and Simpson,A.J.
TITLE Shotgun sequencing of the human transcriptome with ORF expressed sequence tags
JOURNAL Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
MEDLINE 20202663
COMMENT Contact: Simpson A.J.G.
Laboratory of Cancer Genetics
Ludwig Institute for Cancer Research
Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP, Brazil
Tel: +55-11-2704922
Fax: +55-11-2707001
Email: asimpson@ludwig.org.br
This sequence was derived from the FAPESP/LICR Human Cancer Genome Project. This entry can be seen in the following URL (http://www.ludwig.org.br/scripts/gethtml2.pl?tl=ILO&t2=ILO-HN0039-101100-500-c03&t3=2000-11-10&t4=1)
Seq primer: puc 18 forward
High quality sequence stop: 466.
Location/Qualifiers
1. .466
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_lib="HN0039"
/dev_stage="Adult"
/note="Organ: head normal; Vector: puc18; Site_1: SmaI; Site_2: SmaI; A mini-library was made by cloning products derived from ORESTES PCR (U.S. Letters Patent application No. 196,716 - Ludwig Institute for Cancer Research) profiles into the pUC 18 vector. Reverse transcription of tissue mRNA and cDNA amplification were performed under low stringency conditions."
BASE COUNT 98 a 138 c 117 g 113 t
ORIGIN
Alignment Scores:
Pred. No.: 3.5e-29 Length: 466
Score: 327.00 Matches: 57
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 12 Gaps: 0
09-833799-13B (1-57) x BF824935 (1-466)
QY 1 AlaGlnGluProValSerThrLysProGlySerCysProIleLeu 20
Db 184 GCGAAGAGCCAGTCAAGGTCAGTCTCCACTAAGCCTGGCTCCTGCCCATATATCTTG 243

Db 283 GCGCAAGAGCCAGTCAAAGGTCACAGTCCAGTCTCCACTAAGCCTGGCTCCTGCCCATATCTTG 224
QY 21 IleArgCysAlaMetLeuAsnProAsnArgCysLeuLysAspThrAspCysProGly 40
Db 223 ATCCGGTGGCCCATGTTGAATCCCGCTAACCCGCTGCTTGAAGATACTGACTGCCCAGGA 164
QY 41 IleLysLysCysCysGluGlySerCysGlyMetAlaCysPheValProGln 57
Db 163 ATCAAGAAGTGCTGTGAAGGCTCTGCGGGATGGCCTGTTTCGTTCCCCAG 113
RESULT 50
BF837734/c
LOCUS BF837734 471 bp mRNA linear EST 13-JAN-2001
DEFINITION QV3-HT1016-221100-480-g03 HT1016 Homo sapiens cDNA, mRNA sequence.
ACCESSION BF837734
VERSION BF837734.1 GI:12189956
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 471)
Dias Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Briones,M.R.,
Nagai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.F.,
Goldman,G.H., Carvalho,A.F., Matsukuma,A., Baia,G.S., Simpson,D.H.,
Brunstein,A., deOliveira,P.S., Bucher,P., Jongeneel,C.V., O'Hare
,M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and
Simpson,A.J.
Shotgun sequencing of the human transcriptome with ORF expressed
sequence tags
Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
20202663
Contact: Simpson A.J.G.
Laboratory of Cancer Genetics
Ludwig Institute for Cancer Research
Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
Brazil
Tel: +55-11-2704922
Fax: +55-11-2707001
Email: asimpson@ludwig.org.br
This sequence was derived from the FAPESP/LICR Human Cancer Genome
Project. This entry can be seen in the following URL
(http://www.ludwig.org.br/scripts/gethtml2.pl?tl=QV3&t2=QV3-HT1016-
221100-480-g03&t3=2000-11-22&t4=1)
Seq primer: puc 18 forward
High quality sequence stop: 468.
Location/Qualifiers
1. .471
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_lib="HT1016"
/dev_stage="Adult"
/note="Organ: head neck; Vector: puc18; Site 1: SmaI;
Site 2: SmaI; A mini-library was made by cloning products
derived from ORESTES PCR (U.S. Letters Patent application
No. 196,716 - Ludwig Institute for Cancer Research)
profiles into the pUC 18 vector. Reverse transcription of
tissue mRNA and cDNA amplification were performed under
low stringency conditions."
BASE COUNT 116 a 114 c 138 g 102 t 1 others
ORIGIN
Alignment Scores:
Pred. No.: 3.54e-29 Length: 471
Score: 327.00 Matches: 57
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 12 Gaps: 0
09-833799-13B (1-57) x BF837734 (1-471)

QY 1 AlaGlnGluProValLysGlyProValSerThrLysProGlySerCysProIleLeu 20
Db 299 GCGCAAGAGCCAGTCAAAGGTCACAGTCTCCACTAAGCCTGGCTCCTGCCCATATCTTG 240
QY 21 IleArgCysAlaMetLeuAsnProAsnArgCysLeuLysAspThrAspCysProGly 40
Db 239 ATCCGGTGGCCCATGTTGAATCCCGCTAACCCGCTGCTTGAAGATACTGACTGCCCAGGA 180
QY 41 IleLysLysCysCysGluGlySerCysGlyMetAlaCysPheValProGln 57
Db 179 ATCAAGAAGTGCTGTGAAGGCTCTTGGGGATGGCCTGTTTCGTTCCCCAG 129
RESULT 51
BF824932
LOCUS BF824932 477 bp mRNA linear EST 13-JAN-2001
DEFINITION IL0-HN0039-101100-500-a12 HN0039 Homo sapiens cDNA, mRNA sequence.
ACCESSION BF824932
VERSION BF824932.1 GI:12166998
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 477)
Dias Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Briones,M.R.,
Nagai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.F.,
Goldman,G.H., Carvalho,A.F., Matsukuma,A., Baia,G.S., Simpson,D.H.,
Brunstein,A., deOliveira,P.S., Bucher,P., Jongeneel,C.V., O'Hare
,M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and
Simpson,A.J.
Shotgun sequencing of the human transcriptome with ORF expressed
sequence tags
Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
20202663
Contact: Simpson A.J.G.
Laboratory of Cancer Genetics
Ludwig Institute for Cancer Research
Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
Brazil
Tel: +55-11-2704922
Fax: +55-11-2707001
Email: asimpson@ludwig.org.br
This sequence was derived from the FAPESP/LICR Human Cancer Genome
Project. This entry can be seen in the following URL
(http://www.ludwig.org.br/scripts/gethtml2.pl?tl=IL0&t2=IL0-HN0039-
101100-500-a12&t3=2000-11-10&t4=1)
Seq primer: puc 18 forward
High quality sequence stop: 476.
Location/Qualifiers
1. .477
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_lib="HN0039"
/dev_stage="Adult"
/note="Organ: head normal; Vector: puc18; Site 1: SmaI;
Site 2: SmaI; A mini-library was made by cloning products
derived from ORESTES PCR (U.S. Letters Patent application
No. 196,716 - Ludwig Institute for Cancer Research)
profiles into the pUC 18 vector. Reverse transcription of
tissue mRNA and cDNA amplification were performed under
low stringency conditions."
BASE COUNT 103 a 139 c 121 g 114 t
ORIGIN
Alignment Scores:
Pred. No.: 3.6e-29 Length: 477
Score: 327.00 Matches: 57
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 12 Gaps: 0
09-833799-13B (1-57) x BF824932 (1-477)

QY 1 AlaGlnGluProValLysGlyProValSerThrLysProGlySerCysProIleIleLeu 20
Db 184 GCGCAAGAGCCAGTCAAGGTCAGTCTCCACTAAGCTGGCTCCTGCCCATTTATCTTG 243

QY 21 IleArgCysAlaMetLeuAsnProProAsnArgCysLeuLysAspThrAspCysProGly 40
Db 244 ATCCGGTGGCCCATGTTGAATCCCCCTAACCCGCTGCTTGAAGATACTGACTGCCAGGA 303

QY 41 IleLysLysCysCysGluGlySerCysGlyMetAlaCysPheValProGln 57
Db 304 ATCAAGAAGTCTGTGAAGGCTCTTGGGGATGGCCTGTTTCGTTCCCCAG 354

RESULT 52
AA582866/c
LOCUS
DEFINITION nn72a09.s1 NCI CGAP Lar1 Homo sapiens cDNA clone IMAGE:1089400 3'
similar to gb:58717.rnal ELAFIN PRECURSOR (HUMAN); contains element
HGR repetitive element ; mRNA sequence.

ACCESSION AA582866
VERSION AA582866.1 GI:2360226
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE 1 (bases 1 to 480)
AUTHORS NCI-CGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.
TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
JOURNAL Unpublished (1997)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: L. Jeffrey Medeiros, M.D., Michael R.
Emmert-Buck, M.D., Ph.D.
cDNA Library Preparation: Stratagene, Inc.
cDNA Library Arrayed by: Greg Lennon, Ph.D.
DNA Sequencing by: Washington University Genome Sequencing Center
Clone distribution: NCI-CGAP clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
www-bio.llnl.gov/bbrp/image/image.html
Insert length: 647 Std Error: 0.00
Seq primer: -40ml3 fwd. ET from Amersham
High quality sequence stop: 430.
Location/Qualifiers
1. .480
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:1089400"
/clone_lib="NCI CGAP Lar1"
/tissue_type="larynx"
/lab_host="SOLR (kanamycin resistant)"
/note="Organ: larynx; Vector: Bluescript SK-; Site 1:
EcoRI; Site 2: XhoI; Cloned unidirectionally. Primer:
Oligo dT. Larynx. 5' adaptor sequence: 5' GAATTCGGCAGG
3' 3' adaptor sequence: 5' CTCGACTTTTCTTTTCTTTTCTTTT 3'
Average insert size: 0.9 kb."

BASE COUNT 113 a 107 c 144 g 116 t
ORIGIN

Alignment Scores:
Pred. No.: 3.62e-29 Length: 480
Score: 327.00 Matches: 57
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 9 Gaps: 0

09-833799-13B (1-57) x AA582866 (1-480)

QY 1 AlaGlnGluProValLysGlyProValSerThrLysProGlySerCysProIleIleLeu 20
Db 363 GCGCAAGAGCCAGTCAAGGTCAGTCTCCACTAAGCTGGTCTGCCCATTTATCTTG 304

QY 21 IleArgCysAlaMetLeuAsnProProAsnArgCysLeuLysAspThrAspCysProGly 40
Db 303 ATCCGGTGGCCCATGTTGAATCCCCCTAACCCGCTGCTTGAAGATACTGACTGCCAGGA 244

QY 41 IleLysLysCysCysGluGlySerCysGlyMetAlaCysPheValProGln 57
Db 243 ATCAAGAAGTCTGTGAAGGCTCTTGGGGATGGCCTGTTTCGTTCCCCAG 193

RESULT 53
BF825204
LOCUS
DEFINITION IL0-HN0039-131100-501-a07 HN0039 Homo sapiens cDNA, mRNA sequence.
ACCESSION BF825204
VERSION BF825204.1 GI:12167445
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE 1 (bases 1 to 487)
AUTHORS Dias Neto, E., Garcia Correa, R., Verjovski-Almeida, S., Briones, M.R.,
Nagai, M.A., da Silva, W. Jr., Zago, M.A., Bordin, S., Costa, F.F.,
Goldman, G.H., Carvalho, A.F., Matsukuma, A., Baia, G.S., Simpson, D.H.,
Brunstein, A., deOliveira, P.S., Bucher, P., Jongeneel, C.V., O'Hare,
M.J., Soares, F., Brentani, R.R., Reis, L.F., de Souza, S.J. and
Simpson, A.J.
TITLE Shotgun sequencing of the human transcriptome with ORF expressed
sequence tags
JOURNAL Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
MEDLINE 20202663
COMMENT Contact: Simpson A.J.G.
Laboratory of Cancer Genetics
Ludwig Institute for Cancer Research
Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
Brazil
Tel: +55-11-2704922
Fax: +55-11-2707001
Email: asimpson@ludwig.org.br
This sequence was derived from the FAPESP/LICR Human Cancer Genome
Project. This entry can be seen in the following URL
(<http://www.ludwig.org.br/scripts/gethtml2.pl?ti=IL0&t2=IL0-HN0039-131100-501-a07&t3=2000-11-13&t4=1>)
Seq primer: puc 18 forward
High quality sequence start: 11
High quality sequence stop: 487.
Location/Qualifiers
1. .487
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_lib="HN0039"
/dev_stage="Adult"
/note="Organ: head normal; Vector: puc18; Site 1: SmaI;
Site 2: SmaI; A mini-library was made by cloning products
derived from ORESTES PCR (U.S. Letters Patent application
No. 196,716 - Ludwig Institute for Cancer Research)
profiles into the pUC 18 vector. Reverse transcription of
tissue mRNA and cDNA amplification were performed under
low stringency conditions."

BASE COUNT 100 a 146 c 120 g 121 t
ORIGIN

Alignment Scores:
Pred. No.: 3.69e-29 Length: 487
Score: 327.00 Matches: 57
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 12 Gaps: 0

09-833799-13B (1-57) x BF825204 (1-487)

QY 1 AlaGlnGluProValLysGlyProValSerThrLysProGlySerCysProIleIleLeu 20

Db 205 GCGCAAGAGCCAGTCAAAGGTCAGTCTCCACTAAGCCTGGCTCCTGCCCATATCTTG 264
Qy 21 IleArgCysAlaMetLeuAsnProAsnArgCysLeuLysAspThrAspCysProGly 40
Db 265 ATCCGGTGGCCCATGCTGAATCCCCCTAACCGCTGCTTGAAGAATACTGACTGCCAGGA 324
Qy 41 IleLysLysCysCysGluGlySerCysGlyMetAlaCysPheValProGln 57
Db 325 ATCAAGAAGTCTGTGAAGGCTCTTGGGGATGGCCTGTTTCGTTCCCCAG 375
RESULT 54
BF002099/c
LOCUS BF002099 490 bp mRNA linear EST 06-OCT-2000
DEFINITION 7g99b04.x1 NCI CGAP Col6 Homo sapiens cDNA clone IMAGE:3314575 3'
similar to SW:ELAF HUMAN P19957 ELAFIN PRECURSOR ;contains element
MER28 repetitive element ;, mRNA sequence.
ACCESSION BF002099
VERSION BF002099.1 GI:10702374
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
AUTHORS 1 (bases 1 to 490)
TITLE NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
NATIONAL Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
JOURNAL Unpublished (1997)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: Ilan Kirsch, M.D., Michael R. Emmert-Buck, M.D.,
Ph.D.
CDNA Library Preparation: M. Bento Soares, Ph.D.
CDNA Library Arrayed by: Greg Lennon, Ph.D.
DNA Sequencing by: Washington University Genome Sequencing Center
Clone distribution: NCI-CGAP clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL, send email to:
info@image.llnl.gov
Seq primer: -40UP from Gibco
High quality sequence stop: 421.
Location/Qualifiers
1. .490
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:3314575"
/clone_lib="NCI CGAP Col6"
/tissue_type="colon tumor, RER+"
/lab_host="DH10B"
/note="Organ: colon; Vector: pT7T3D-Pac (Pharmacia) with a
modified polylinker; Site 1: Not I; Site 2: Eco RI;
Plasmid DNA from the normalized library NCI CGAP Col6 was
prepared, and ss circles were made in vitro. Following HAP
purification, this DNA was used as tracer in a subtractive
hybridization reaction. The driver was PCR-amplified cDNAs
from a pool of 5,000 clones made from the same library
(cloneIDs 1057416-1061255, and 1144584-1145351).
Subtraction by Bento Soares and M. Fatima Bonaldo. "
BASE COUNT 119 a 113 c 146 g 112 t
ORIGIN
Alignment Scores:
Pred. No.: 3.71e-29 Length: 490
Score: 327.00 Matches: 57
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 12 Gaps: 0
09-833799-13B (1-57) x BF002099 (1-490)
Qy 1 AlaGlnGluProValLysGlyProValSerThrLysProGlySerCysProIleLeu 20

Db 355 GCGCAAGAGCCAGTCAAAGGTCAGTCTCCACTAAGCCTGGCTCCTGCCCATATCTTG 296
Qy 21 IleArgCysAlaMetLeuAsnProAsnArgCysLeuLysAspThrAspCysProGly 40
Db 295 ATCCGGTGGCCCATGTTGAATCCCCCTAACCGCTGCTTGAAGAATACTGACTGCCAGGA 236
Qy 41 IleLysLysCysCysGluGlySerCysGlyMetAlaCysPheValProGln 57
Db 235 ATCAAGAAGTCTGTGAAGGCTCTTGGGGATGGCCTGTTTCGTTCCCCAG 185
RESULT 55
BF837735
LOCUS BF837735 498 bp mRNA linear EST 13-JAN-2001
DEFINITION QV3-HT1016-221100-480-g08 HT1016 Homo sapiens cDNA, mRNA sequence.
ACCESSION BF837735
VERSION BF837735.1 GI:12189957
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
AUTHORS 1 (bases 1 to 498)
Dias Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Briones,M.R.,
Nagai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.F.,
Goldman,G.H., Carvalho,A.F., Matsukuma,A., Baia,G.S., Simpson,D.H.,
Brunstein,A., deOliveira,P.S., Bucher,P., Jongeneel,C.V., O'Hare
M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and
Simpson,A.J.
TITLE Shotgun sequencing of the human transcriptome with ORF expressed
sequence tags
JOURNAL Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
MEDLINE 20202663
COMMENT Contact: Simpson A.J.G.
Laboratory of Cancer Genetics
Ludwig Institute for Cancer Research
Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
Brazil
Tel: +55-11-2704922
Fax: +55-11-2707001
Email: asimpson@ludwig.org.br
This sequence was derived from the FAPESP/LICR Human Cancer Genome
Project. This entry can be seen in the following URL
(http://www.ludwig.org.br/scripts/gethtml2.pl?t1=QV3&t2=QV3-HT1016-
221100-480-g08&t3=2000-11-22&t4=1)
Seq primer: puc 18 forward
High quality sequence stop: 498.
Location/Qualifiers
1. .498
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_lib="HT1016"
/dev_stage="Adult"
/note="Organ: head neck; Vector: puc18; Site 1: SmaI;
Site 2: SmaI; A mini-library was made by cloning products
derived from ORESTES PCR (U.S. Letters Patent application
No. 196,716 - Ludwig Institute for Cancer Research)
profiles into the pUC 18 vector. Reverse transcription of
tissue mRNA and cDNA amplification were performed under
low stringency conditions."
BASE COUNT 104 a 146 c 128 g 120 t
ORIGIN
Alignment Scores:
Pred. No.: 3.78e-29 Length: 498
Score: 327.00 Matches: 57
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 12 Gaps: 0
09-833799-13B (1-57) x BF837735 (1-498)
Qy 1 AlaGlnGluProValLysGlyProValSerThrLysProGlySerCysProIleLeu 20

Db 203 GCGAAGAGCCAGTCAAGGTCAGTCTCCACTAAGCCTGGCTCTGCTCCATTTATCTTG 262
QY 21 IleArgCysAlaMetLeuAsnProProAsnArgCysLeuLysAspThrAspCysProGly 40
Db 263 ATCCGGTGGCCATGTTGAATCCCCCTAACCGTCTTGGGGATGGCCTGTTTCGTTCCCCAG 322
QY 41 IleLysLysCysCysGluGlySerCysGlyMetAlaCysPheValProGln 57
Db 323 ATCAAGAAGTCTGTGAAGGCTCTTGGGGATGGCCTGTTTCGTTCCCCAG 373
RESULT 56
BQ941085 522 bp mRNA linear EST 21-AUG-2002
LOCUS AGENCOURT_8794612 NIH_MGC_18 Homo sapiens cDNA clone IMAGE:6374832
DEFINITION 5', mRNA sequence.
ACCESSION BQ941085
VERSION BQ941085.1 GI:22356563
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 522)
NIH-MGC http://mgc.nci.nih.gov/.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: DCTD/DTP/Gazdar
cDNA Library Preparation: Rubin Laboratory
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLCM2554 row: h column: 01
High quality sequence stop: 521.
Features Location/Qualifiers
1..522
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:6374832"
/clone_lib="NIH_MGC_18"
/tissue_type="large cell carcinoma"
/lab_host="DH10B (phage-resistant)"
/note="Organ: lung; Vector: pOTB7; Site 1: XhoI; Site 2:
EcoRI; cDNA made by oligo-dT priming. Directionally cloned
into EcoRI/XhoI sites using the following 5' adaptor:
GGCAGCAG(G). Library constructed by Ling Hong in the
laboratory of Gerald M. Rubin (University of California,
Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and
Superscript II RT (Life Technologies). Note: this is a
NIH_MGC Library."
BASE COUNT 124 a 155 c 121 g 122 t
ORIGIN
Alignment Scores: 4e-29 Length: 522
Pred. No.: 327.00 Matches: 57
Score: 100.00% Conservative: 0
Percent Similarity: 100.00% Mismatches: 0
Best Local Similarity: 100.00% Indels: 0
Query Match: 14 Gaps: 0
DB: 14
09-833799-13B (1-57) x BQ941085 (1-522)
QY 1 AlaGlnGluProValLysGlyProValSerThrLysProGlySerCysProIleIleLeu 20
Db 156 GCGAAGAGCCAGTCAAGGTCAGTCTCCACTAAGCCTGGCTCTGCTCCATTTATCTTG 215
QY 21 IleArgCysAlaMetLeuAsnProProAsnArgCysLeuLysAspThrAspCysProGly 40

Db 216 ATCCGGTGGCCATGTTGAATCCCCCTAACCGCTGCTTGAAGATAGTACTGACTGCCCCAGGA 275
QY 41 IleLysLysCysCysGluGlySerCysGlyMetAlaCysPheValProGln 57
Db 276 ATCAAGAAGTCTGTGAAGGCTCTTGGGGATGGCCTGTTTCGTTCCCCAG 326
RESULT 57
BQ490472 565 bp mRNA linear EST 27-MAR-2001
LOCUS 602519610F1 NIH_MGC_18 Homo sapiens cDNA clone IMAGE:4637906 5',
DEFINITION mRNA sequence.
ACCESSION BQ490472
VERSION BQ490472.1 GI:13451982
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 565)
NIH-MGC http://mgc.nci.nih.gov/.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: DCTD/DTP/Gazdar
cDNA Library Preparation: Ling Hong/Rubin Laboratory
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLCM1399 row: d column: 03
High quality sequence stop: 564.
Features Location/Qualifiers
1..565
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:4637906"
/clone_lib="NIH_MGC_18"
/tissue_type="large cell carcinoma"
/lab_host="DH10B (phage-resistant)"
/note="Organ: lung; Vector: pOTB7; Site 1: XhoI; Site 2:
EcoRI; cDNA made by oligo-dT priming. Directionally cloned
into EcoRI/XhoI sites using the following 5' adaptor:
GGCAGCAG(G). Library constructed by Ling Hong in the
laboratory of Gerald M. Rubin (University of California,
Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and
Superscript II RT (Life Technologies). Note: this is a
NIH_MGC Library."
BASE COUNT 141 a 163 c 130 g 131 t
ORIGIN
Alignment Scores: 4.39e-29 Length: 565
Pred. No.: 327.00 Matches: 57
Score: 100.00% Conservative: 0
Percent Similarity: 100.00% Mismatches: 0
Best Local Similarity: 100.00% Indels: 0
Query Match: 12 Gaps: 0
DB: 12
09-833799-13B (1-57) x BQ490472 (1-565)
QY 1 AlaGlnGluProValLysGlyProValSerThrLysProGlySerCysProIleIleLeu 20
Db 188 GCGAAGAGCCAGTCAAGGTCAGTCTCCACTAAGCCTGGCTCTGCTCCATTTATCTTG 247
QY 21 IleArgCysAlaMetLeuAsnProProAsnArgCysLeuLysAspThrAspCysProGly 40
Db 248 ATCCGGTGGCCATGTTGAATCCCCCTAACCGCTGCTTGAAGATAGTACTGACTGCCCCAGGA 307
QY 41 IleLysLysCysCysGluGlySerCysGlyMetAlaCysPheValProGln 57
Db 308 ATCAAGAAGTCTGTGAAGGCTCTTGGGGATGGCCTGTTTCGTTCCCCAG 358

```
RESULT 58
BE772161/c
LOCUS
DEFINITION CM4-FT0104-230600-215-c09 FT0104 Homo sapiens cDNA, mRNA sequence.
ACCESSION BE772161
VERSION BE772161.1 GI:10225819
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 361)
AUTHORS Dias Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Briones,M.R.,
Nagai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.F.,
Goldman,G.H., Carvalho,A.F., Matsukuma,A., Baia,G.S., Simpson,D.H.,
Brunstein,A., deOliveira,P.S., Bucher,P., Jongeneel,C.V., O'Hare
,M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and
Simpson,A.J.
TITLE Shotgun sequencing of the human transcriptome with ORF expressed
sequence tags
JOURNAL Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
MEDLINE 20202663
COMMENT Contact: Simpson A.J.G.
Laboratory of Cancer Genetics
Ludwig Institute for Cancer Research
Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
Brazil
Tel: +55-11-2704922
Fax: +55-11-2707001
Email: asimpson@ludwig.org.br
This sequence was derived from the FAPESP/LICR Human Cancer Genome
Project. This entry can be seen in the following URL
(http://www.ludwig.org.br/scripts/gethtml2.pl?tl=&t2=CM4-FT0104-230
600-215-c09&t3=2000-06-23&t4=1)
Seq primer: puc 18 forward
High quality sequence start: 25
High quality sequence stop: 360.
FEATURES
Location/Qualifiers
1..361
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_lib="FT0104"
/dev_stage="Adult"
/note="Organ: prostate tumor; Vector: puc18; Site 1: SmaI;
Site 2: SmaI; A mini-library was made by cloning products
derived from ORESTES PCR (U.S. Letters Patent application
No. 196,716 - Ludwig Institute for Cancer Research)
profiles into the pUC 18 vector. Reverse transcription of
tissue mRNA and cDNA amplification were performed under
low stringency conditions."
BASE COUNT 86 a 85 c 105 g 85 t
ORIGIN
Alignment Scores:
Pred. No.: 5,89e-29 Length: 361
Score: 324.00 Matches: 56
Percent Similarity: 100.00% Conservative: 1
Best Local Similarity: 98.25% Mismatches: 0
Query Match: 99.08% Indels: 0
DB: 12 Gaps: 0
09-833799-13B (1-57) x BE772161 (1-361)
QY 1 AlaGlnGluProValSerThrLysProGlySerCysProIleLeu 20
|||||:|||||
Db 256 GCGAAGATCCAGTCAAAGTCCAGTCTCCACTAAGCTGGCTCCTGCCCATATCTTG 197
GCGCAAGAGCCAGTCAAAGTCCAGTCTCCACTAAGCTGGCTCCTGCCCATATCTTG
QY 21 IleArgCysAlaMetLeuAsnProProAsnArgCysLeuLysAspThrAspCysProGly 40
|||||:|||||
Db 196 ATCCGGTGGCCATGTTGAATCCCCCTAACCGCTGCTTGAAAGATACTGACTGCCCAGGA 137
ATCCGGTGGCCATGTTGAATCCCCCTAACCGCTGCTTGAAAGATACTGACTGCCCAGGA
QY 41 IleLysLysCysCysGluGlySerCysGlyMetAlaCysPheValProGln 57
|||||:|||||
```

```
Db 136 ATCAAGAAGTGCTGTGAAGGCTCTTTCGGGATGGCTGTTTCGTTCCCCAG 86
|||||:|||||
RESULT 59
BI061069
LOCUS
DEFINITION IL3-UT0116-020201-464-F04 UT0116 Homo sapiens cDNA, mRNA sequence.
ACCESSION BI061069
VERSION BI061069.1 GI:14468596
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 421)
AUTHORS Dias Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Briones,M.R.,
Nagai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.F.,
Goldman,G.H., Carvalho,A.F., Matsukuma,A., Baia,G.S., Simpson,D.H.,
Brunstein,A., deOliveira,P.S., Bucher,P., Jongeneel,C.V., O'Hare
,M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and
Simpson,A.J.
TITLE Shotgun sequencing of the human transcriptome with ORF expressed
sequence tags
JOURNAL Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
MEDLINE 20202663
COMMENT Contact: Simpson A.J.G.
Laboratory of Cancer Genetics
Ludwig Institute for Cancer Research
Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
Brazil
Tel: +55-11-2704922
Fax: +55-11-2707001
Email: asimpson@ludwig.org.br
This sequence was derived from the FAPESP/LICR Human Cancer Genome
Project. This entry can be seen in the following URL
(http://www.ludwig.org.br/scripts/gethtml2.pl?tl=IL3&t2=IL3-UT0116-
020201-464-F04&t3=2001-02-02&t4=1)
Seq primer: puc 18 forward
High quality sequence stop: 353.
FEATURES
Location/Qualifiers
1..421
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_lib="UT0116"
/dev_stage="Adult"
/note="Organ: uterus_tumor; Vector: puc18; Site 1: SmaI;
Site 2: SmaI; A mini-library was made by cloning products
derived from ORESTES PCR (U.S. Letters Patent application
No. 196,716 - Ludwig Institute for Cancer Research)
profiles into the pUC 18 vector. Reverse transcription of
tissue mRNA and cDNA amplification were performed under
low stringency conditions."
BASE COUNT 94 a 116 c 113 g 98 t
ORIGIN
Alignment Scores:
Pred. No.: 7,06e-29 Length: 421
Score: 324.00 Matches: 56
Percent Similarity: 100.00% Conservative: 1
Best Local Similarity: 98.25% Mismatches: 0
Query Match: 99.08% Indels: 0
DB: 13 Gaps: 0
09-833799-13B (1-57) x BI061069 (1-421)
QY 1 AlaGlnGluProValSerThrLysProGlySerCysProIleLeu 20
|||||:|||||
Db 185 GCGAAGAGCCAGTCAAAGTCCAGTCTCCACTAAGCTGGCTCCTGCCCATATCTTG 244
GCGCAAGAGCCAGTCAAAGTCCAGTCTCCACTAAGCTGGCTCCTGCCCATATCTTG
QY 21 IleArgCysAlaMetLeuAsnProProAsnArgCysLeuLysAspThrAspCysProGly 40
|||||:|||||
Db 245 ATCCGGTGGCCATGTTGAATCCCCCTAACCGCTGCTTGAAAGATACTGACTGCCCAGGA 304
ATCCGGTGGCCATGTTGAATCCCCCTAACCGCTGCTTGAAAGATACTGACTGCCCAGGA
```

QY 41 ILeLysLysCysCysGluGlySerCysGlyMetAlaCysPheValProGln 57
||||:|||||
Db 305 ATCAGGAAGTGTGTGAAGGCTCTTCGGGATGGCCTGTTCTGTTCCCCAG 355
RESULT 60
BE711100
LOCUS BE711100 263 bp mRNA linear EST 12-SEP-2000
DEFINITION RC3-HT0649-270700-012-f03 HT0649 Homo sapiens cDNA, mRNA sequence.
ACCESSION BE711100
VERSION BE711100.1 GI:10099365
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 263)
AUTHORS Dias Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Briones,M.R.,
Nagai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.F.,
Goldman,G.H., Carvalho,A.F., Matsukuma,A., Baia,G.S., Simpson,D.H.,
Brunstein,A., deOliveira,P.S., Bucher,P., Jongeneel,C.V., O'Hare
,M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and
Simpson,A.J.
TITLE Shotgun sequencing of the human transcriptome with ORF expressed
sequence tags
JOURNAL Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
MEDLINE 20202663
COMMENT Contact: Simpson A.J.G.
Laboratory of Cancer Genetics
Ludwig Institute for Cancer Research
Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
Brazil
Tel: +55-11-2704922
Fax: +55-11-2707001
Email: asimpson@ludwig.org.br
This sequence was derived from the FAPESP/LICR Human Cancer Genome
Project. This entry can be seen in the following URL
(http://www.ludwig.org.br/scripts/gethtml2.pl?t1=&t2=RC3-HT0649-270
700-012-f03&t3=2000-07-27&t4=1)
Seq primer: puc 18 forward
High quality sequence start: 50
High quality sequence stop: 263.
FEATURES
source
1..263
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_lib="HT0649"
/dev_stage="Adult"
/note="Organ: head neck; Vector: puc18; Site 1: SmaI;
Site 2: SmaI; A mini-library was made by cloning products
derived from ORESTES PCR (U.S. Letters Patent application
No. 196,716 - Ludwig Institute for Cancer Research)
profiles into the pUC 18 vector. Reverse transcription of
tissue mRNA and cDNA amplification were performed under
low stringency conditions."
BASE COUNT 62 a 71 c 74 g 56 t
ORIGIN
Alignment Scores:
Pred. No.: 5.33e-29 Length: 263
Score: 323.00 Matches: 56
Percent Similarity: 100.00% Conservative: 1
Best Local Similarity: 98.25% Mismatches: 0
Query Match: 98.78% Indels: 0
DB: 12 Gaps: 0
09-833799-13B (1-57) x BE711100 (1-263)
QY 1 AlaGlnGluProValLysGlyProValSerThrLysProGlySerCysProIleLeu 20
|||||
Db 57 GCGCAAGAGCCAGTCAAAGGTCCTCCACTGAGCCTGGCTCCTGCCCATATCTTG 116
QY 21 IleArgCysAlaMetLeuAsnProProAsnArgCysLeuLysAspThrAspCysProGly 40
|||||

Db 117 ATCCGGTCGCCATGTTGAATCCCCCTAACCGCTGCTTGAAGATACTGACTGCCAGGA 176
QY 41 ILeLysLysCysCysGluGlySerCysGlyMetAlaCysPheValProGln 57
|||||
Db 177 ATCAAGAAGTGTGTGAAGGCTCTTCGGGATGGCCTGTTTCGTACCCAG 227
RESULT 61
AI924155/c
LOCUS AI924155 420 bp mRNA linear EST 08-MAR-2000
DEFINITION wn64c08.x1 NCI CGAP Lul9 Homo sapiens cDNA clone IMAGE:2450222 3'
similar to gb:S58717_rnal ELAFIN PRECURSOR (HUMAN); contains element
PTR5 repetitive element ;, mRNA sequence.
ACCESSION AI924155
VERSION AI924155.1 GI:5660119
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 420)
AUTHORS NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
JOURNAL Unpublished (1997)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: Christopher Moskaluk, M.D., Ph.D., Michael R.
Emmert-Buck, M.D., Ph.D.
cDNA Library Preparation: M. Bento Soares, Ph.D.
cDNA Library Arrayed by: Greg Lennon, Ph.D.
DNA Sequencing by: Washington University Genome Sequencing Center
Clone distribution: NCI-CGAP clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
www-bio.llnl.gov/bbrp/image/image.html
Insert Length: 511 Std Error: 0.00
Seq primer: -40UP from Gibco.
FEATURES
source
1..420
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:2450222"
/clone_lib="NCI CGAP Lul9"
/tissue_type="squamous cell carcinoma, poorly
differentiated (4 pooled tumors, including primary and
metastatic)"
/dev_stage="adult"
/lab_host="DH10B (phage-resistant)"
/note="Organ: lung; Vector: pT7T3D-Pac (Pharmacia) with a
modified polylinker; 1st strand cDNA was prepared from
pooled lung tumor tissue, and was then primed with a Not I
- oligo(dT) primer. Double-stranded cDNA was ligated to
Eco RI adaptors (Pharmacia), digested with Not I and
cloned into the Not I and Eco RI sites of the modified
pT7T3 vector. Library went through one round of
normalization. Library constructed by Bento Soares and M.
Fatima Bonaldo."
BASE COUNT 102 a 89 c 130 g 99 t
ORIGIN
Alignment Scores:
Pred. No.: 9.26e-29 Length: 420
Score: 323.00 Matches: 56
Percent Similarity: 100.00% Conservative: 1
Best Local Similarity: 98.25% Mismatches: 0
Query Match: 98.78% Indels: 0
DB: 9 Gaps: 0
09-833799-13B (1-57) x AI924155 (1-420)
QY 1 AlaGlnGluProValLysGlyProValSerThrLysProGlySerCysProIleLeu 20
|||||
Db 358 GCGCAAGAGCCAGTCAAAGGTCCTCCACTAAGCCTGGCTCCTGCCCATATCTTG 299

QY 21 IleArgCysAlaMetLeuAsnProProAsnArgCysLeuLysAspThrAspCysProGly 40
|||||
Db 298 ATCCGGTGGCCCATGTTGAATCCCCCTAACCGCTGCTTGAAAGATACTGACTGCCCAGGA 239
|||||
QY 41 IleLysLysCysCysGluGlySerCysGlyMetAlaCysPheValProGln 57
|||||
Db 238 ATCAAGAAGTGCTGTAAAGGCTCTTGGGGAATGGCTGTTTCGTTCCCCAG 188
|||||
RESULT 62
BF835679 447 bp mRNA linear EST 13-JAN-2001
LOCUS QV3-HT1016-151100-463-c10 HT1016 Homo sapiens cDNA, mRNA sequence.
DEFINITION BF835679
ACCESSION BF835679
VERSION BF835679.1 GI:12186947
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 447)
AUTHORS Dias Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Briones,M.R.,
Nagai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.F.,
Goldman,G.H., Carvalho,A.F., Matsukuma,A., Baia,G.S., Simpson,D.H.,
Brunstein,A., deOliveira,P.S., Bucher,P., Jongeneel,C.V., O'Hare
,M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and
Simpson,A.J.
TITLE Shotgun sequencing of the human transcriptome with ORF expressed
sequence tags
JOURNAL Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
MEDLINE 20202663
COMMENT Contact: Simpson A.J.G.
Laboratory of Cancer Genetics
Ludwig Institute for Cancer Research
Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
Brazil
Tel: +55-11-2704922
Fax: +55-11-2707001
Email: asimpson@ludwig.org.br
This sequence was derived from the FAPESP/LICR Human Cancer Genome
Project. This entry can be seen in the following URL
(http://www.ludwig.org.br/scripts/gethtml2.pl?t1=QV3&t2=QV3-HT1016-
151100-463-c10&t3=2000-11-15&t4=1)
Seq primer: puc 18 forward
High quality sequence start: 8
High quality sequence stop: 447.
FEATURES
source
1..447
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_lib="HT1016"
/dev_stage="Adult"
/note="Organ: head_neck; Vector: puc18; Site_1: SmaI;
Site_2: SmaI; A mini-library was made by cloning products
derived from ORESTES PCR (U.S. Letters Patent application
No. 196,716 - Ludwig Institute for Cancer Research)
profiles into the pUC 18 vector. Reverse transcription of
tissue mRNA and cDNA amplification were performed under
low stringency conditions."
BASE COUNT 96 a 128 c 117 g 106 t
ORIGIN
Alignment Scores:
Pred. No.: 9.97e-29 Length: 447
Score: 323.00 Matches: 56
Percent Similarity: 100.00% Conservative: 1
Best Local Similarity: 98.25% Mismatches: 0
Query Match: 98.78% Indels: 0
DB: 12 Gaps: 0
09-833799-13B (1-57) x BF835679 (1-447)
QY 1 AlaGlnGluProValLysGlyProValSerThrLysProGlySerCysProIleLeu 20
|||||

Db 189 GCGCGAGAGCCAGTCAAAGGTCCAGTCTCCACTAAGCCTGGCTCCTGCCCATTTATCTTG 248
QY 21 IleArgCysAlaMetLeuAsnProProAsnArgCysLeuLysAspThrAspCysProGly 40
|||||
Db 249 ATCCGGTGGCCCATGTTGAATCCCCCTAACCGCTGCTTGAAAGATACTGACTGCCCAGGA 308
|||||
QY 41 IleLysLysCysCysGluGlySerCysGlyMetAlaCysPheValProGln 57
|||||
Db 309 ATCAAGAAGTGCTGTGAAGGCTCTTGGCGGATGGCTGTTTCGTTCCCCAG 359
|||||
RESULT 63
AW845140/c 453 bp mRNA linear EST 19-MAY-2000
LOCUS IL0-CT0008-140599-007 CT0008 Homo sapiens cDNA, mRNA sequence.
DEFINITION AW845140
ACCESSION AW845140
VERSION AW845140.1 GI:7940657
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 453)
AUTHORS Dias Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Briones,M.R.,
Nagai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.F.,
Goldman,G.H., Carvalho,A.F., Matsukuma,A., Baia,G.S., Simpson,D.H.,
Brunstein,A., deOliveira,P.S., Bucher,P., Jongeneel,C.V., O'Hare
,M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and
Simpson,A.J.
TITLE Shotgun sequencing of the human transcriptome with ORF expressed
sequence tags
JOURNAL Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
MEDLINE 20202663
COMMENT Contact: Simpson A.J.G.
Laboratory of Cancer Genetics
Ludwig Institute for Cancer Research
Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
Brazil
Tel: +55-11-2704922
Fax: +55-11-2707001
Email: asimpson@ludwig.org.br
This sequence was derived from the FAPESP/LICR Human Cancer Genome
Project. This entry can be seen in the following URL
(http://www.ludwig.org.br/scripts/gethtml2.pl?t1=IL0-CT0008-140
599-007&t3=1999-05-14&t4=1)
Seq primer: puc 18 forward
High quality sequence start: 39
High quality sequence stop: 454.
FEATURES
source
1..453
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_lib="CT0008"
/dev_stage="Adult"
/note="Organ: colon; Vector: puc18; Site_1: SmaI; Site_2:
SmaI; A mini-library was made by cloning products derived
from ORESTES PCR (U.S. Letters Patent application No. 196
,716 - Ludwig Institute for Cancer Research)
profiles into the pUC 18 vector. Reverse transcription of tissue
mRNA and cDNA amplification were performed under low
stringency conditions."
BASE COUNT 106 a 110 c 139 g 98 t
ORIGIN
Alignment Scores:
Pred. No.: 1.01e-28 Length: 453
Score: 323.00 Matches: 56
Percent Similarity: 98.25% Conservative: 0
Best Local Similarity: 98.25% Mismatches: 1
Query Match: 98.78% Indels: 0
DB: 10 Gaps: 0
09-833799-13B (1-57) x AW845140 (1-453)

Qy 1 AlaGlnGluProValLysGlyProValSerThrLysProGlySerCysProIleLeu 20
|||||
Db 307 GCGCAAGAGCCAGTCAAAGGTCAGCTCCACTAAGCTGGCTCCTGCCCATATCTTG 248
|||||
Qy 21 IleArgCysAlaMetLeuAsnProProAsnArgCysLeuLysAspThrAspCysProGly 40
|||||
Db 247 ATCCGGTGGCCCATGTTGAATCCCTAACCGCTGCTTGAAGAGATACTGACTGCCAGGA 188
|||||
Qy 41 IleLysLysCysCysGluGlySerCysGlyMetAlaCysPheValProGln 57
|||||
Db 187 ATCAAGAAGTGCTGTGAAGGCTCTTGGGGATGGCCTGTTTCGTTCCCCAG 137
|||||
RESULT 64
AW845142/c
LOCUS AW845142 455 bp mRNA linear EST 19-MAY-2000
DEFINITION IL0-CT0008-140599-001 CT0008 Homo sapiens cDNA, mRNA sequence.
ACCESSION AW845142
VERSION AW845142.1 GI:7940659
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 455)
AUTHORS Dias Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Briones,M.R.,
Nagai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.F.,
Goldman,G.H., Carvalho,A.F., Matsukuma,A., Baia,G.S., Simpson,D.H.,
Brunstein,A., deOliveira,P.S., Bucher,P., Jongeneel,C.V., O'Hare
,M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and
Simpson,A.J.
TITLE Shotgun sequencing of the human transcriptome with ORF expressed
sequence tags
JOURNAL Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
MEDLINE 20202663
COMMENT Contact: Simpson A.J.G.
Laboratory of Cancer Genetics
Ludwig Institute for Cancer Research
Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
Brazil
Tel: +55-11-2704922
Fax: +55-11-2707001
Email: asimpson@ludwig.org.br
This sequence was derived from the FAPESP/LICR Human Cancer Genome
Project. This entry can be seen in the following URL
(http://www.ludwig.org.br/scripts/gethtml2.pl?t1=et2=IL0-CT0008-140
599-001&t3=1999-05-14&t4=1)
Seq primer: puc 18 forward
High quality sequence stop: 455.
Location/Qualifiers
FEATURES
source
1. .455
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_lib="CT0008"
/dev_stage="Adult"
/note="Organ: colon; Vector: puc18; Site_1: SmaI; Site_2:
SmaI; A mini-library was made by cloning products derived
from ORESTES PCR (U.S. Letters Patent application No. 196
,716 - Ludwig Institute for Cancer Research) profiles
into the pUC 18 vector. Reverse transcription of tissue
mRNA and cDNA amplification were performed under low
stringency conditions."

BASE COUNT 108 a 116 c 134 g 97 t
ORIGIN
Alignment Scores:
Pred. No.: 1.02e-28 Length: 455
Score: 323.00 Matches: 56
Percent Similarity: 100.00% Conservative: 1
Best Local Similarity: 98.25% Mismatches: 0
Query Match: 98.78% Indels: 0
DB: 10 Gaps: 0
09-833799-13B (1-57) x AW845142 (1-455)

Qy 1 AlaGlnGluProValLysGlyProValSerThrLysProGlySerCysProIleLeu 20
|||||
Db 282 GCGCAAGAGCCAGTCAAAGGTCAGCTCCACTAAGCTGGCTCCTGCCCATATCTTG 223
|||||
Qy 21 IleArgCysAlaMetLeuAsnProProAsnArgCysLeuLysAspThrAspCysProGly 40
|||||
Db 222 ATCCGGTGGCCCATGTTGAATCCCTAACCGCTGCTTGAAGAGATACTGACTGCCAGGA 163
|||||
Qy 41 IleLysLysCysCysGluGlySerCysGlyMetAlaCysPheValProGln 57
|||||
Db 162 ATCAAGAAGTGCTGTGAAGGCTCTTGGGGATGGCCTGTTTCGTTCCCCGG 112
|||||
RESULT 65
AI858070/c
LOCUS AI858070 444 bp mRNA linear EST 21-DEC-1999
DEFINITION wj70a03.x1 NCI CGAP Lul9 Homo sapiens cDNA clone IMAGE:2408140 3'
similar to gb:S58717_rnal ELAFIN PRECURSOR (HUMAN);contains element
THR repetitive element ;, mRNA sequence.
ACCESSION AI858070
VERSION AI858070.1 GI:5511686
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 444)
AUTHORS NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
JOURNAL Unpublished (1997)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: Christopher Moskaluk, M.D., Ph.D., Michael R.
Emmert-Buck, M.D., Ph.D.
cDNA Library Preparation: M. Bento Soares, Ph.D.
cDNA Library Arrayed by: Greg Lennon, Ph.D.
DNA Sequencing by: Washington University Genome Sequencing Center
Clone distribution: NCI-CGAP clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
www-bio.llnl.gov/bbrp/image/image.html
Insert Length: 551 Std Error: 0.00
Seq primer: -40UP from Gibco.
Location/Qualifiers
FEATURES
source
1. .444
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_lib="IMAGE:2408140"
/clone_lib="NCI CGAP_Lul9"
/tissue_type="squamous cell carcinoma, poorly
differentiated (4 pooled tumors, including primary and
metastatic)"
/dev_stage="adult"
/lab_host="DH10B (phage-resistant)"
/note="Organ: lung; Vector: pT7T3D-Pac (Pharmacia) with a
modified polylinker; 1st strand cDNA was prepared from
pooled lung tumor tissue, and was then primed with a Not I
- oligo(dT) primer. Double-stranded cDNA was ligated to
Eco RI adaptors (Pharmacia), digested with Not I and
cloned into the Not I and Eco RI sites of the modified
pT7T3 vector. Library went through one round of
normalization. Library constructed by Bento Soares and M.
Fatima Bonaldo."

BASE COUNT 108 a 96 c 134 g 105 t 1 others
ORIGIN
Alignment Scores:
Pred. No.: 1.3e-28 Length: 444
Score: 322.00 Matches: 56
Percent Similarity: 98.25% Conservative: 0
Best Local Similarity: 98.25% Mismatches: 1
Query Match: 98.47% Indels: 0
DB: 9 Gaps: 0

Percent Similarity: 98.25% Conservative: 0
Best Local Similarity: 98.25% Mismatches: 1
Query Match: 98.17% Indels: 0
DB: 12 Gaps: 0

09-833799-13B (1-57) x BF836201 (1-374)

QY 1 AlaGlnGluProValLysGlyProValSerThrLysProGlySerCysProIleLeu 20
|||||
Db 169 GCGAAGAGCCAGTCAAAGGTCCAGTCTCCACTAAGCTGGCTCTGCCCATTTCTTG 228
|||||
QY 21 IleArgCysAlaMetLeuAsnProProAsnArgCysLeuLysAspThrAspCysProGly 40
|||||
Db 229 ATCCGTCGCGCCATGTTGAATCCCCCTAAGCGCTGCTTGAAGATATACTGCTGCTCCAGGA 288
|||||
QY 41 IleLysLysCysCysGluGlySerCysGlyMetAlaCysPheValProGln 57
|||||
Db 289 ATCAAGAAGTCTGTGAAGGCTCTTGGGGATGGCCTGTTTCGTTCCCCAG 339
|||||

RESULT 68

BF825196

LOCUS BF825196 501 bp mRNA linear EST 13-JAN-2001
DEFINITION IL0-HN0039-131100-501-d10 HN0039 Homo sapiens cDNA, mRNA sequence.

ACCESSION BF825196

VERSION BF825196.1 GI:12167431

KEYWORDS EST.

SOURCE human.

ORGANISM

Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE

AUTHORS

Dias Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Briones,M.R.,
Nagai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.F.,
Goldman,G.H., Carvalho,A.F., Matsukuma,A., Baia,G.S., Simpson,D.H.,
Brunstein,A., deOliveira,P.S., Bucher,P., Jongeneel,C.V., O'Hare
,M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and
Simpson,A.J.

Shotgun sequencing of the human transcriptome with ORF expressed

TITLE

sequence tags
Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)

JOURNAL

MEDLINE

COMMENT

Contact: Simpson A.J.G.
Laboratory of Cancer Genetics
Ludwig Institute for Cancer Research
Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
Brazil

Tel: +55-11-2704922

Fax: +55-11-2707001

Email: asimpson@ludwig.org.br

This sequence was derived from the FAPESP/LICR Human Cancer Genome
Project. This entry can be seen in the following URL

(http://www.ludwig.org.br/scripts/gethtml2.pl?tl=IL0&t2=IL0-HN0039-
131100-501-d10&t3=2000-11-13&t4=1)

Seq primer: puc 18 forward

High quality sequence stop: 501.

FEATURES

source

1. 501
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_lib="HN0039"
/dev_stage="Adult"
/note="Organ: head normal; Vector: puc18; Site 1: Smal;
Site 2: Smal; A mini-library was made by cloning products
derived from ORESTES PCR (U.S. Letters Patent application
No. 196,716 - Ludwig Institute for Cancer Research)
profiles into the pUC 18 vector. Reverse transcription of
tissue mRNA and cDNA amplification were performed under
low stringency conditions."
109 a 146 c 127 g 119 t

BASE COUNT

ORIGIN

Alignment Scores: 1.97e-28 Length: 501

Pred. No.: 7

Score: 321.00 Matches: 56
Percent Similarity: 98.25% Conservative: 0
Best Local Similarity: 98.25% Mismatches: 1
Query Match: 98.17% Indels: 0
DB: 12 Gaps: 0

09-833799-13B (1-57) x BF825196 (1-501)

QY 1 AlaGlnGluProValLysGlyProValSerThrLysProGlySerCysProIleLeu 20
|||||
Db 202 GCGAAGAGCCAGTCAAAGGTCCAGTCTCCACTAAGCTGGCTCTGCCCATTTCTTG 261
|||||

QY 21 IleArgCysAlaMetLeuAsnProProAsnArgCysLeuLysAspThrAspCysProGly 40
|||||
Db 262 ATCCGTCGCGCCATGTTGAATCCCCCTAAGCGCTGCTTGAAGATATACTGCTGCTCCAGGA 321
|||||

QY 41 IleLysLysCysCysGluGlySerCysGlyMetAlaCysPheValProGln 57
|||||
Db 322 ATCAAGAAGTCTGTGAAGGCTCTTGGGGATGGACTGTTTCGTACCCAG 372
|||||

RESULT 69

BE711083/c

LOCUS BE711083 316 bp mRNA linear EST 12-SEP-2000
DEFINITION RC3-HT0649-270700-012-f06 HT0649 Homo sapiens cDNA, mRNA sequence.

ACCESSION BE711083

VERSION BE711083.1 GI:10099348

KEYWORDS EST.

SOURCE human.

ORGANISM

Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE

AUTHORS

Dias Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Briones,M.R.,
Nagai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.F.,
Goldman,G.H., Carvalho,A.F., Matsukuma,A., Baia,G.S., Simpson,D.H.,
Brunstein,A., deOliveira,P.S., Bucher,P., Jongeneel,C.V., O'Hare
,M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and
Simpson,A.J.

Shotgun sequencing of the human transcriptome with ORF expressed

sequence tags

Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)

JOURNAL

MEDLINE

COMMENT

Contact: Simpson A.J.G.
Laboratory of Cancer Genetics
Ludwig Institute for Cancer Research

Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
Brazil

Tel: +55-11-2704922

Fax: +55-11-2707001

Email: asimpson@ludwig.org.br

This sequence was derived from the FAPESP/LICR Human Cancer Genome
Project. This entry can be seen in the following URL

(http://www.ludwig.org.br/scripts/gethtml2.pl?tl=&t2=RC3-HT0649-270
700-012-f06&t3=2000-07-27&t4=1)

Seq primer: puc 18 forward

High quality sequence start: 22

High quality sequence stop: 316.

FEATURES

source

1. 316
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_lib="HT0649"
/dev_stage="Adult"
/note="Organ: head neck; Vector: puc18; Site 1: Smal;
Site 2: Smal; A mini-library was made by cloning products
derived from ORESTES PCR (U.S. Letters Patent application
No. 196,716 - Ludwig Institute for Cancer Research)
profiles into the pUC 18 vector. Reverse transcription of
tissue mRNA and cDNA amplification were performed under
low stringency conditions."
68 a 73 c 114 g 61 t

BASE COUNT

ORIGIN

Alignment Scores:

Pred. No.: 1.51e-28 Length: 316
Score: 320.00 Matches: 55
Percent Similarity: 100.00% Conservative: 2
Best Local Similarity: 96.49% Mismatches: 0
Query Match: 97.86% Indels: 0
DB: 12 Gaps: 0

09-833799-13B (1-57) x BE7111083 (1-316)

QY 1 AlaGlnGluProValSerThrLysProGlySerCysProIleLeu 20
|||
Db 289 GCGCAAGAGCCACTCAAGGTCAGTCTCCACTAAGCCTGGCTCTGCCCCATTATCTTG 230
|||
QY 21 IleArgCysAlaMetLeuAsnProProAsnArgCysLeuLysAspThrAspCysProGly 40
|||
Db 229 ATCCGGTGGCCCATGTTGAATCCCCCTAACCGCTGCTTGAAGATACTGACTGCCCAGGA 170
|||
QY 41 IleLysLysCysCysGluGlySerCysGlyMetAlaCysPheValProGln 57
|||
Db 169 ATCAAGAAGTGCTGTGAAGGCTCTTGGGGATGGCCCTGTTTCGTTCCCGGT 119
|||

RESULT 70
AW844969

LOCUS AW844969 344 bp mRNA linear EST 19-MAY-2000
DEFINITION MR0-CT0006-280599-008 CT0006 Homo sapiens cDNA, mRNA sequence.
ACCESSION AW844969
VERSION AW844969.1 GI:7940577
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
AUTHORS 1 (bases 1 to 344)
Dias Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Briones,M.R.,
Nagai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.F.,
Goldman,G.H., Carvalho,A.F., Matsukuma,A., Baia,G.S., Simpson,D.H.,
Brunstein,A., deOliveira,P.S., Bucher,P., Jongeneel,C.V., O'Hare
,M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and
Simpson,A.J.

TITLE Shotgun sequencing of the human transcriptome with ORF expressed
sequence tags
JOURNAL Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
MEDLINE 20202663
COMMENT Contact: Simpson A.J.G.
Laboratory of Cancer Genetics
Ludwig Institute for Cancer Research
Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
Brazil
Tel: +55-11-2704922
Fax: +55-11-2707001
Email: asimpson@ludwig.org.br

This sequence was derived from the FAPESP/LICR Human Cancer Genome
Project. This entry can be seen in the following URL
(http://www.ludwig.org.br/scripts/gethtml2.pl?tl=&t2=MR0-CT0006-280
599-008&t3=1999-05-28&t4=1)
Seq primer: puc 18 forward
High quality sequence start: 7
High quality sequence stop: 345.
Location/Qualifiers
1. 344
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_lib="CT0006"
/dev_stage="Adult"
/note="Organ: colon; Vector: puc18; Site 1: SmaI; Site 2:
SmaI; A mini-library was made by cloning products derived
from ORESTES PCR (U.S. Letters Patent application No. 196
,716 - Ludwig Institute for Cancer Research) profiles
into the pUC 18 vector. Reverse transcription of tissue
mRNA and cDNA amplification were performed under low
stringency conditions."

FEATURES

source
1. 344
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_lib="CT0006"
/dev_stage="Adult"
/note="Organ: colon; Vector: puc18; Site 1: SmaI; Site 2:
SmaI; A mini-library was made by cloning products derived
from ORESTES PCR (U.S. Letters Patent application No. 196
,716 - Ludwig Institute for Cancer Research) profiles
into the pUC 18 vector. Reverse transcription of tissue
mRNA and cDNA amplification were performed under low
stringency conditions."

BASE COUNT 82 a 89 c 88 g 85 t

ORIGIN

Alignment Scores:
Pred. No.: 1.66e-28 Length: 344
Score: 320.00 Matches: 56
Percent Similarity: 98.25% Conservative: 0
Best Local Similarity: 98.25% Mismatches: 1
Query Match: 97.86% Indels: 0
DB: 10 Gaps: 0

09-833799-13B (1-57) x AW844969 (1-344)

QY 1 AlaGlnGluProValSerThrLysProGlySerCysProIleLeu 20
|||
Db 174 GCGCAGGAGCCAGTCAAAGGTCCAGTCTCCACTAAGCCTGGCTCTGCCCATATCTTG 233
|||
QY 21 IleArgCysAlaMetLeuAsnProProAsnArgCysLeuLysAspThrAspCysProGly 40
|||
Db 234 ATCCGGTGGCCCATGTTGAATCCCCCTAACCGCTGCTTGAAGATACTGACTGCCCAGGA 293
|||
QY 41 IleLysLysCysCysGluGlySerCysGlyMetAlaCysPheValProGln 57
|||
Db 294 ATCAAGAAGTGCTGTGAAGGCTCTTGGGATATGGCCTGTTTCGTTCCCCAG 344
|||

RESULT 71
BM768573

LOCUS BM768573 428 bp mRNA linear EST 04-MAR-2002
DEFINITION K-EST0051573 S14K402 Homo sapiens cDNA clone S14K402-6-G01 5', mRNA
sequence.
ACCESSION BM768573
VERSION BM768573.1 GI:19098188
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
AUTHORS 1 (bases 1 to 428)
Kim,N.S., Hahn,Y., Oh,J.H., Lee,J.Y., Ahn,H.Y., Chu,M.Y., Kim,M.R.,
Oh,K.J., Cheong,J.E., Sohn,H.Y., Kim,J.M., Park,H.S., Kim,S. and
Kim,Y.S.

TITLE 21C Frontier Korean EST Project 2001
JOURNAL Unpublished (2002)
COMMENT Contact: Kim YS
Genome Research Center
Korea Research Institute of Bioscience & Biotechnology
52 Boeun-dong Yuseong-gu, Daejeon 305-333, South Korea
Tel: +82-42-860-4470
Fax: +82-42-860-4409
Email: yongsung@mail.kribb.re.kr
Plate: 6 row: G column: 01
High quality sequence stop: 428.

FEATURES

source
1. 428
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_lib="S14K402-6-G01"
/clone_lib="S14K402"
/cell_line="K402"
/lab_host="Top10F"
/note="Organ: Stomach; Vector: pTZ18RP1; Site 1: EcoRI;
Site 2: NotI; The poly (A)+ RNA was dephosphorylated with
bacterial alkaline phosphatase (BAP) and then decapped
with tobacco acid pyrophosphatase (TAP). The decapped
intact mRNA was ligated with DNA-RNA linker including EcoR
I site by treatment of T4 RNA ligase and the first strand
cDNA was synthesized from oligo dT-selected mRNA by
priming with dT-tailed vector. The dT-tailed vector was
adjusted to have about 60nt. The cDNA vector was
circularized with E. coli DNA ligase after digestion of
EcoRI which site is also included in vector. An RNA strand
converted to a DNA strand by Okayama-Berg method. The
obtained cDNA vectors were used for transfection of
competent cells E. coli Top10F' by electroporation method.

The cDNA libraries constructed by this method are full-length enriched cDNA library."

BASE COUNT
ORIGIN

101 a 120 c 108 g 99 t

Alignment Scores:

Pred. No.: 2.15e-28 Length: 428
Score: 320.00 Matches: 56
Percent Similarity: 98.25% Conservative: 0
Best Local Similarity: 98.25% Mismatches: 1
Query Match: 97.86% Indels: 0
DB: 14 Gaps: 0

09-833799-13B (1-57) x BM768573 (1-428)

QY 1 AlaGlnGluProValLysGlyProValSerThrLysProGlySerCysProIleIleLeu 20
|||||
Db 200 GCGCAAGAGCCAGTCAAAGGTCCAGTCTCCACTAAGCCTGGCTCCTGCCCATATCTTG 259
|||||
QY 21 IleArgCysAlaMetLeuAsnProProAsnArgCysLeuLysAspThrAspCysProGly 40
|||||
Db 260 ATCCCGTGGCCCATGTTGAATCCCCCTAACCGCTGCTTGAAGATACTGACTGCCAGGA 319
|||||
QY 41 IleLysLysCysCysGluGlySerCysGlyMetAlaCysPheValProGln 57
|||||
Db 320 ATCAAGAAATGCTGTGAAGGCTCTTGGGGATGGCCTGTTCTCCCCAG 370
|||||

RESULT 72

BI021805/c

LOCUS BI021805 394 bp mRNA linear EST 14-JUN-2001
DEFINITION CM3-MT0316-170101-673-e02 MT0316 Homo sapiens cDNA, mRNA sequence.

ACCESSION BI021805

VERSION BI021805.1 GI:14428435

KEYWORDS EST.

SOURCE human.

ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 394)

REFERENCE

AUTHORS

Dias Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Briones,M.R.,
Nagai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.F.,
Goldman,G.H., Carvalho,A.F., Matsukuma,A., Baia,G.S., Simpson,D.H.,
Brunstein,A., deOliveira,P.S., Bucher,P., Jongeneel,C.V., O'Hare
M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and
Simpson,A.J.

TITLE Shotgun sequencing of the human transcriptome with ORF expressed
sequence tags

JOURNAL Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)

MEDLINE 20202663

COMMENT

Contact: Simpson A.J.G.

Laboratory of Cancer Genetics

Ludwig Institute for Cancer Research

Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
Brazil

Tel: +55-11-2704922

Fax: +55-11-2707001

Email: asimpson@ludwig.org.br

This sequence was derived from the FAPESP/LICR Human Cancer Genome
Project. This entry can be seen in the following URL

(http://www.ludwig.org.br/scripts/gethtml2.pl?t1=CM3&t2=CM3-MT0316-

170101-673-e02&t3=2001-01-17&t4=1)

Seq primer: puc 18 forward

High quality sequence stop: 394.

Location/Qualifiers

FEATURES

source

1. .394

/organism="Homo sapiens"

/db_xref="taxon:9606"

/clone_lib="MT0316"

/dev_stage="Adult"

/note="Organ: marrow; Vector: puc18; Site_1: SmaI; Site_2:
SmaI; A mini-library was made by cloning products derived
from ORESTES PCR (U.S. Letters Patent application No. 196
,716 - Ludwig Institute for Cancer Research) profiles

into the pUC 18 vector. Reverse transcription of tissue
mRNA and cDNA amplification were performed under low
stringency conditions."

BASE COUNT 91 a 106 c 108 g 89 t
ORIGIN

Alignment Scores:

Pred. No.: 2.57e-28 Length: 394
Score: 319.00 Matches: 56
Percent Similarity: 98.25% Conservative: 0
Best Local Similarity: 98.25% Mismatches: 1
Query Match: 97.55% Indels: 0
DB: 13 Gaps: 0

09-833799-13B (1-57) x BI021805 (1-394)

QY 1 AlaGlnGluProValLysGlyProValSerThrLysProGlySerCysProIleIleLeu 20
|||||
Db 229 GCGCAAGAGGCAGTCAAAGGTCCAGTCTCCACTAAGCCTGGCTCCTGCCCATATCTTG 170
|||||

QY 21 IleArgCysAlaMetLeuAsnProProAsnArgCysLeuLysAspThrAspCysProGly 40
|||||

Db 169 ATCCGCTGGCCCATGTTGAATCCCCCTAACCGCTGCTTGAAGATACTGACTGCCAGGA 110
|||||

QY 41 IleLysLysCysCysGluGlySerCysGlyMetAlaCysPheValProGln 57
|||||

Db 109 ATCAAGAAATGCTGTGAAGGCTCTTGGGGATGGCCTGTTCTCCCCAG 59
|||||

RESULT 73

AA586983/c

LOCUS AA586983

DEFINITION nm9e10.s1 NCI CGAP Lar1 Homo sapiens cDNA clone IMAGE:1089162 3'

similar to gb:S58717_rna1 ELAFIN PRECURSOR (HUMAN); contains element
THR repetitive element ; mRNA sequence.

ACCESSION AA586983

VERSION AA586983.1 GI:2397797

KEYWORDS EST.

SOURCE human.

ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 386)

REFERENCE

AUTHORS

NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.

National Cancer Institute, Cancer Genome Anatomy Project (CGAP),

Tumor Gene Index

Unpublished (1997)

Contact: Robert Strausberg, Ph.D.

Email: cgapbs-r@mail.nih.gov

Tissue Procurement: L. Jeffrey Medeiros, M.D., Michael R.

Emmert-Buck, M.D., Ph.D.

cDNA Library Preparation: Stratagene, Inc.

cDNA Library Arrayed by: Greg Lennon, Ph.D.

DNA Sequencing by: Washington University Genome Sequencing Center

Clone distribution: NCI-CGAP clone distribution information can be

found through the I.M.A.G.E. Consortium/LLNL at:

www-bio.llnl.gov/bbrp/image/image.html

Insert Length: 642 Std Error: 0.00

Seq primer: -40m13 fwd. ET from Amersham

High quality sequence stop: 241.

Location/Qualifiers

FEATURES

source

1. .386

/organism="Homo sapiens"

/db_xref="taxon:9606"

/clone="IMAGE:1089162"

/clone_lib="NCI-CGAP_Lar1"

/tissue_type="larynx"

/lab_host="SOLR (kanamycin resistant)"

/note="Organ: larynx; Vector: Bluescript SK-; Site_1:

EcoRI; Site_2: XhoI; Cloned unidirectionally. Primer:

Oligo dT. Larynx. 5' adaptor sequence: 5' GAATTCGGCAG

3' 3' adaptor sequence: 5' CTCGAGTTTTTTTTTTTTTTT 3'

Average insert size: 0.9 kb."

BASE COUNT 86 a 75 c 121 g 104 t

normalization. Library constructed by Bento Soares and M.
Fatima Bonaldo."

BASE COUNT 106 a 91 c 125 g 101 t
ORIGIN

Alignment Scores:
Pred. No.: 3.68e-28 Length: 423
Score: 318.00 Matches: 55
Percent Similarity: 98.25% Conservative: 1
Best Local Similarity: 96.49% Mismatches: 1
Query Match: 97.25% Indels: 0
DB: 9 Gaps: 0

09-833799-13B (1-57) x AI830872 (1-423)

QY 1 AlaGlnGluProValLysGlyProValSerThrLysProGlySerCysProIleIleLeu 20
Db 359 GCGCAAGAGCCAGTCAAAGGTCCAGTCTCCAAATAAGCCTGGCTCCTGCCCAATTATCTTG 300
QY 21 IleArgCysAlaMetLeuAsnProProAsnArgCysLeuLysAspThrAspCysProGly 40
Db 299 ATCCGGTGGCCATGTTGAATCCCCCTAACCGGTGTTTGAAGAAGAACTGACTGCCCAGGA 240
QY 41 IleLysLysCysCysGluGlySerCysGlyMetAlaCysPheValProGln 57
Db 239 ATCAAGAAGTCTGTGAAGGTTCTTGGGGGATGGCCTGTTTCGTTCCCCAG 189

Search completed: February 15, 2003, 21:12:55
Job time : 1467 secs

Sun Feb 16 09:13:26 2003

09-833799-13b.rnpb

5	143	43.7	594	10	US-09-964-824A-582	Sequence 582, App
6	143	43.7	594	10	US-09-954-456-1989	Sequence 1989, Ap
7	143	43.7	594	10	US-09-865-812-1	Sequence 1, Appli
8	112	34.3	762	9	US-09-992-598-344	Sequence 344, App
9	112	34.3	762	9	US-09-989-293A-344	Sequence 344, App
10	112	34.3	762	9	US-09-989-735-344	Sequence 344, App
11	112	34.3	762	9	US-09-990-444-344	Sequence 344, App
12	112	34.3	762	9	US-09-989-730-344	Sequence 344, App
13	112	34.3	762	9	US-09-990-436-344	Sequence 344, App
14	112	34.3	762	9	US-09-991-181-344	Sequence 344, App
15	112	34.3	762	9	US-09-993-687-344	Sequence 344, App
16	112	34.3	762	9	US-09-989-734-344	Sequence 344, App
17	112	34.3	762	9	US-09-997-653-344	Sequence 344, App
18	112	34.3	762	9	US-09-993-667-344	Sequence 344, App
19	112	34.3	762	9	US-09-990-438-344	Sequence 344, App
20	112	34.3	762	9	US-09-990-562-344	Sequence 344, App
21	112	34.3	762	9	US-09-997-428-344	Sequence 344, App
22	112	34.3	762	9	US-09-997-666-344	Sequence 344, App
23	112	34.3	762	10	US-09-989-722-344	Sequence 344, App
24	112	34.3	762	10	US-09-989-723-344	Sequence 344, App
25	112	34.3	762	10	US-09-989-279-344	Sequence 344, App
26	112	34.3	762	10	US-09-989-727-344	Sequence 344, App
27	112	34.3	762	10	US-09-989-731-344	Sequence 344, App
28	112	34.3	762	10	US-09-989-732-344	Sequence 344, App
29	112	34.3	762	10	US-09-991-073-344	Sequence 344, App
30	112	34.3	762	10	US-09-990-442-344	Sequence 344, App
31	112	34.3	762	10	US-09-991-163-344	Sequence 344, App
32	112	34.3	762	10	US-09-993-604-344	Sequence 344, App
33	112	34.3	762	10	US-09-990-456-344	Sequence 344, App
34	112	34.3	762	10	US-09-989-721-344	Sequence 344, App
35	107	32.7	411	10	US-09-964-824A-495	Sequence 495, App
36	107	32.7	411	10	US-09-954-456-1987	Sequence 1987, Ap
37	107	32.7	411	10	US-09-967-768A-3	Sequence 3, Appli
38	105	32.1	292	10	US-09-884-441-377	Sequence 377, App
39	98	30.0	724	9	US-10-152-661-589	Sequence 589, App
40	97	29.7	583	10	US-09-967-768A-214	Sequence 214, App
41	95	29.1	469	10	US-09-864-761-1629	Sequence 1629, Ap
42	94.5	28.9	228	10	US-09-790-264-57	Sequence 57, Appl
43	94.5	28.9	437	10	US-09-823-038A-29	Sequence 29, Appl
44	94.5	28.9	471	10	US-09-790-264-55	Sequence 55, Appl
45	93	28.4	622	10	US-09-925-300-421	Sequence 421, App
46	93	28.4	643	9	US-09-924-340-41	Sequence 41, Appl
47	93	28.4	643	9	US-09-992-600A-41	Sequence 41, Appl
48	93	28.4	783	9	US-10-093-766-57	Sequence 57, Appl
49	92.5	28.3	6314	10	US-09-954-456-319	Sequence 319, App
50	92.5	28.3	6314	10	US-09-954-456-1230	Sequence 1230, Ap
51	85	26.0	753	9	US-09-852-797-34	Sequence 34, Appl
52	85	26.0	753	10	US-09-853-161-34	Sequence 34, Appl
53	85	26.0	753	10	US-09-852-659A-34	Sequence 34, Appl
54	85	26.0	783	9	US-09-852-797-50	Sequence 50, Appl
55	85	26.0	783	10	US-09-853-161-50	Sequence 50, Appl
56	85	26.0	783	10	US-09-852-659A-50	Sequence 50, Appl
57	80.5	24.6	354	10	US-09-852-659A-121	Sequence 121, App
58	78.5	24.0	334	10	US-09-833-381-747	Sequence 747, App
59	74.5	22.8	2082	10	US-09-819-136-1	Sequence 1, Appli
60	73.5	22.5	243	10	US-09-894-882-240	Sequence 240, App
61	73.5	22.5	243	10	US-09-894-882-249	Sequence 249, App
62	71	21.7	3836	12	US-10-044-090-750	Sequence 750, App
63	70.5	21.6	2937	10	US-09-974-300-2288	Sequence 2288, Ap
64	69.5	21.3	753	10	US-09-910-943-661	Sequence 661, App
65	69.5	21.3	465237	10	US-09-933-267A-1	Sequence 1, Appli
66	69	21.1	148	10	US-09-864-761-18387	Sequence 18387, A
67	69	21.1	403	10	US-09-884-441-198	Sequence 198, App
68	68	20.8	511	10	US-09-864-761-12659	Sequence 12659, A
69	67.5	20.6	2108	10	US-09-962-832-225	Sequence 225, App
70	67	20.5	302250	10	US-09-962-832-154	Sequence 154, App
71	66	20.2	2853	10	US-09-815-242-7863	Sequence 7863, Ap
72	66	20.2	4656	10	US-09-891-216-6	Sequence 6, Appli
73	66	20.2	10172	10	US-09-954-456-702	Sequence 702, App
74	66	20.2	10284	9	US-09-974-298-46	Sequence 46, Appl
75	66	20.2	10284	12	US-10-044-090-47	Sequence 47, Appl
76	64.5	19.7	243	10	US-09-894-882-243	Sequence 243, App
77	64.5	19.7	243	10	US-09-894-882-255	Sequence 255, App

OM protein - nucleic search, using frame_plus_p2n model

Run on: February 15, 2003, 20:49:11 ; Search time 108 Seconds
(without alignments)
268.810 Million cell updates/sec

Title: 09-833799-13B
Perfect score: 327
Sequence: 1 aqepvkgpvstkgpcpiil.....cpgikkcccgscgmactvpq 57

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 424239 seqs, 254661826 residues

Total number of hits satisfying chosen parameters: 848478

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 100 summaries

Command line parameters:
-MODEL=frame+p2n.model -DEV=xlp
-Q=/cgn2_1/USPTO_spool/US09833799/runat_12022003_120044_23813/app_query.fasta_1.199
-DB=Published Applications NA -QFMT=fastap -SUFFIX=rnpb -MINMATCH=0.1
-LOOPCL=0 -LOOPEXT=0 -UNITS=bits -START=1 -END=-1 -MATRIX=blosum62
-TRANS=human40.cdi -LIST=100 -DOCLIGN=200 -THR SCORE=pct -THR MAX=100
-THR MIN=0 -ALIGN=75 -MODE=LOCAL -OUTFMT=pto -NORM=ext -HEAPSIZE=500 -MINLEN=0
-MAXLEN=200000000 -USER=US09833799 @CGN 1 1 80 @runat_12022003_120044_23813
-NCPU=6 -ICPU=3 -NO XLPXY -NO MMAP -LARGEQUERY -NEG SCORES=0 -WAIT -LONGLOG
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=10 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Published Applications NA:
1: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq:
2: /cgn2_6/ptodata/1/pubpna/PTC_NEW_PUB.seq:
3: /cgn2_6/ptodata/1/pubpna/US05_NEW_PUB.seq:
4: /cgn2_6/ptodata/1/pubpna/US06_PUBCOMB.seq:
5: /cgn2_6/ptodata/1/pubpna/US07_NEW_PUB.seq:
6: /cgn2_6/ptodata/1/pubpna/PTCUS_PUBCOMB.seq:
7: /cgn2_6/ptodata/1/pubpna/US08_NEW_PUB.seq:
8: /cgn2_6/ptodata/1/pubpna/US08_PUBCOMB.seq:
9: /cgn2_6/ptodata/1/pubpna/US09_NEW_PUB.seq:
10: /cgn2_6/ptodata/1/pubpna/US09_PUBCOMB.seq:
11: /cgn2_6/ptodata/1/pubpna/US10_NEW_PUB.seq:
12: /cgn2_6/ptodata/1/pubpna/US10_PUBCOMB.seq:
13: /cgn2_6/ptodata/1/pubpna/US60_NEW_PUB.seq:
14: /cgn2_6/ptodata/1/pubpna/US60_PUBCOMB.seq:

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length DB	ID	Description
C 1	327	100.0	480	10	US-09-920-300A-1393
C 2	327	100.0	480	12	US-10-033-528-1393
C 3	327	100.0	2309	10	US-09-954-456-1126
4	327	100.0	2309	10	US-09-954-456-1798

78	64.5	19.7	243	10	US-09-894-882-349
79	64.5	19.7	764	10	US-09-770-445-942
80	64.5	19.7	1168	10	US-09-772-179-1
C 81	64.5	19.7	1320	10	US-09-815-242-9218
C 82	64.5	19.7	1364	10	US-09-967-768A-324
C 83	64.5	19.7	2657	9	US-10-108-605-156
C 84	64.5	19.7	2687	9	US-10-108-605-78
85	64	19.6	2793	9	US-10-007-271-3
86	64	19.6	3282	9	US-10-007-271-1
87	64	19.6	4513	9	US-09-989-442-163
88	64	19.6	4513	10	US-09-764-853-921
89	64	19.6	4531	9	US-09-989-442-164
90	64	19.6	4531	10	US-09-764-853-922
91	64	19.6	32177	10	US-09-764-877-3251
92	64	19.6	32207	10	US-09-764-877-3250
93	63.5	19.4	852	10	US-09-764-898-50
C 94	63.5	19.4	852	10	US-09-764-898-123
95	63.5	19.4	1359	9	US-10-000-512-13
96	63	19.3	1644	10	US-09-819-136-7
C 97	62.5	19.1	1417	10	US-09-917-800A-1606
98	62.5	19.1	3396	10	US-09-749-728B-32
99	62.5	19.1	4015	9	US-09-843-676-224
100	62.5	19.1	4015	9	US-09-953-052-1

Sequence 349, App
Sequence 942, App
Sequence 1, Appli
Sequence 9218, Ap
Sequence 324, App
Sequence 156, App
Sequence 78, Appli
Sequence 3, Appli
Sequence 1, Appli
Sequence 163, App
Sequence 921, App
Sequence 164, App
Sequence 922, App
Sequence 3251, Ap
Sequence 3250, Ap
Sequence 50, Appl
Sequence 123, App
Sequence 13, Appl
Sequence 7, Appli
Sequence 1606, Ap
Sequence 32, Appl
Sequence 224, App
Sequence 1, Appli

ALIGNMENTS

RESULT 1
US-09-920-300A-1393/c
; Sequence 1393, Application US/09920300A
; Patent No. US20020136728A1
; GENERAL INFORMATION:
; APPLICANT: King, Gordon E.
; APPLICANT: Meagher, Madeleine Joy
; APPLICANT: Xu, Jiangchun
; APPLICANT: Secrist, Heather
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; TITLE OF INVENTION: AND DIAGNOSIS OF COLON CANCER
; FILE REFERENCE: 210121.547
; CURRENT APPLICATION NUMBER: US/09/920,300A
; NUMBER OF SEQ ID NOS: 1789
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 1393
; LENGTH: 480
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: 27, 56, 343
; OTHER INFORMATION: n = A,T,C or G
US-09-920-300A-1393

Alignment Scores:
Pred. No.: 8.6e-33
Score: 327.00
Length: 480
Percent Similarity: 100.00%
Matches: 57
Best Local Similarity: 100.00%
Conservative: 0
Query Match: 100.00%
Mismatch: 0
DB: 10
Indels: 0
Gaps: 0

09-833799-13B (1-57) x US-09-920-300A-1393 (1-480)
QY 1 AlaGlnGluProValLysGlyProValSerThrLysProGlySerCysProIleLeu 20
Db 369 GCGCAAGAGCCAGTCAAGGTCAGTCCACTAAGCCTGGCTCTGCCCCATTATCTTG 310
QY 21 IleArgCysAlaMetLeuAsnProProAsnArgCysLeuLysAspThrAspCysProGly 40
Db 309 ATCCGGTGGCCATGTTGAATCCCCCTAACCCTGCTTGAAGATACTGACTGCCAGGA 250
QY 41 IleLysLysCysCysGluGlySerCysGlyMetAlaCysPheValProGln 57

Db 249 ATCAAGAAGTGCTGTGAAGGCTCTTGGGGATGGCCTGTTTCGTTCCCCAG 199
RESULT 2
US-10-033-528-1393/c
; Sequence 1393, Application US/10033528
; Patent No. US20020131971A1
; GENERAL INFORMATION:
; APPLICANT: King, Gordon E.
; APPLICANT: Meagher, Madeleine Joy
; APPLICANT: Xu, Jiangchun
; APPLICANT: Secrist, Heather
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; TITLE OF INVENTION: AND DIAGNOSIS OF COLON CANCER
; FILE REFERENCE: 210121.547C1
; CURRENT APPLICATION NUMBER: US/10/033,528
; CURRENT FILING DATE: 2001-12-26
; NUMBER OF SEQ ID NOS: 1896
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 1393
; LENGTH: 480
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: 27, 56, 343
; OTHER INFORMATION: n = A,T,C or G
US-10-033-528-1393

Alignment Scores:
Pred. No.: 8.6e-33
Score: 327.00
Length: 480
Percent Similarity: 100.00%
Matches: 57
Best Local Similarity: 100.00%
Conservative: 0
Query Match: 100.00%
Mismatch: 0
DB: 12
Indels: 0
Gaps: 0

09-833799-13B (1-57) x US-10-033-528-1393 (1-480)
QY 1 AlaGlnGluProValLysGlyProValSerThrLysProGlySerCysProIleLeu 20
Db 369 GCGCAAGAGCCAGTCAAGGTCAGTCCACTAAGCCTGGCTCTGCCCCATTATCTTG 310
QY 21 IleArgCysAlaMetLeuAsnProProAsnArgCysLeuLysAspThrAspCysProGly 40
Db 309 ATCCGGTGGCCATGTTGAATCCCCCTAACCCTGCTTGAAGATACTGACTGCCAGGA 250
QY 41 IleLysLysCysCysGluGlySerCysGlyMetAlaCysPheValProGln 57
Db 249 ATCAAGAAGTGCTGTGAAGGCTCTTGGGGATGGCCTGTTTCGTTCCCCAG 199

RESULT 3
US-09-954-456-1126
; Sequence 1126, Application US/09954456
; Patent No. US20020115057A1
; GENERAL INFORMATION:
; APPLICANT: Young, Paul
; TITLE OF INVENTION: Process for Identifying Anti-Cancer Therapeutic Agents Using Ca
; TITLE OF INVENTION: Sets
; FILE REFERENCE: 689290-76
; CURRENT APPLICATION NUMBER: US/09/954,456
; CURRENT FILING DATE: 2001-09-18
; PRIOR APPLICATION NUMBER: US/60/233,617
; PRIOR FILING DATE: 2000-09-18
; PRIOR APPLICATION NUMBER: US/60/234,052
; PRIOR FILING DATE: 2000-09-20
; PRIOR APPLICATION NUMBER: US/60/234,923
; PRIOR FILING DATE: 2000-09-25
; PRIOR APPLICATION NUMBER: US/60/235,134
; PRIOR FILING DATE: 2000-09-25
; PRIOR APPLICATION NUMBER: US/60/235,637
; PRIOR FILING DATE: 2000-09-26
; PRIOR APPLICATION NUMBER: US/60/235,638
; PRIOR FILING DATE: 2000-09-26

Db 298 GGCCAATGTTTGATGCTTAACCCCAATTTCTGTGAGATGGCCAGTGAACGGT 357

QY 41 IleLysLysCysCysGluGlySerCysGlyMetAlaCysPheValPro 56

Db 358 GACTTGAAGTGTTCATGGGCATGTGTGGAAATCCTGCGTTTCCCT 405

RESULT 6

US-09-954-456-1989

; Sequence 1989, Application US/09954456

; Patent No. US20020115057A1

; GENERAL INFORMATION:

; APPLICANT: Young, Paul

; TITLE OF INVENTION: Process for Identifying Anti-Cancer Therapeutic Agents Using Can

; FILE OF INVENTION: Sets

; FILE REFERENCE: 689290-76

; CURRENT APPLICATION NUMBER: US/09/954,456

; PRIOR FILING DATE: 2001-09-18

; PRIOR APPLICATION NUMBER: US/60/233,617

; PRIOR FILING DATE: 2000-09-18

; PRIOR APPLICATION NUMBER: US/60/234,052

; PRIOR FILING DATE: 2000-09-20

; PRIOR APPLICATION NUMBER: US/60/234,923

; PRIOR FILING DATE: 2000-09-25

; PRIOR APPLICATION NUMBER: US/60/235,134

; PRIOR FILING DATE: 2000-09-25

; PRIOR APPLICATION NUMBER: US/60/235,637

; PRIOR FILING DATE: 2000-09-26

; PRIOR APPLICATION NUMBER: US/60/235,638

; PRIOR FILING DATE: 2000-09-26

; PRIOR APPLICATION NUMBER: US/60/235,711

; PRIOR FILING DATE: 2000-09-27

; PRIOR APPLICATION NUMBER: US/60/235,720

; PRIOR FILING DATE: 2000-09-27

; PRIOR APPLICATION NUMBER: US/60/235,840

; PRIOR FILING DATE: 2000-09-27

; PRIOR APPLICATION NUMBER: US/60/235,863

; NUMBER OF SEQ ID NOS: 2276

; SOFTWARE: PatentIn version 3.0

; SEQ ID NO 1989

; LENGTH: 594

; TYPE: DNA

; ORGANISM: Homo sapiens

US-09-954-456-1989

Alignment Scores:

Pred. No.: 1.7e-09

Score: 143.00

Percent Similarity: 55.36%

Best Local Similarity: 48.21%

Query Match: 43.73%

DB: 10

09-833799-13B (1-57) x US-09-954-456-1989 (1-594)

QY 3 GluProValLysGlyProValSerThr-----LysProGlySerCysProIleIleLeu 20

Db 238 GATCCTGTTGACACCCCAACCCCAAGGAGGAGCCTGGGAAGTGGCCAGTGACTTAT 297

QY 21 IleArgCysAlaMetLeuAsnProAsnArgCysLeuLysAspThrAspCysProGly 40

Db 298 GGCCAATGTTGATGCTTAACCCCAATTTCTGTGAGATGGCCAGTGAACGGT 357

QY 41 IleLysLysCysCysGluGlySerCysGlyMetAlaCysPheValPro 56

Db 358 GACTTGAAGTGTTCATGGGCATGTGTGGAAATCCTGCGTTTCCCT 405

RESULT 7

US-09-865-812-1

; Sequence 1, Application US/09865812

; Patent No. US20020115626A1

; GENERAL INFORMATION:

; APPLICANT: Rastelli, Luca

Alignment Scores:

Pred. No.: 1.7e-09

Score: 143.00

Percent Similarity: 55.36%

Best Local Similarity: 48.21%

Query Match: 43.73%

DB: 10

09-833799-13B (1-57) x US-09-954-456-1989 (1-594)

QY 3 GluProValLysGlyProValSerThr-----LysProGlySerCysProIleIleLeu 20

Db 238 GATCCTGTTGACACCCCAACCCCAAGGAGGAGCCTGGGAAGTGGCCAGTGACTTAT 297

QY 21 IleArgCysAlaMetLeuAsnProAsnArgCysLeuLysAspThrAspCysProGly 40

Db 298 GGCCAATGTTGATGCTTAACCCCAATTTCTGTGAGATGGCCAGTGAACGGT 357

QY 41 IleLysLysCysCysGluGlySerCysGlyMetAlaCysPheValPro 56

Db 358 GACTTGAAGTGTTCATGGGCATGTGTGGAAATCCTGCGTTTCCCT 405

RESULT 8

US-09-992-598-344

; Sequence 344, Application US/09992598

; Patent No. US20020160384A1

; GENERAL INFORMATION:

; APPLICANT: Ashkenazi, Avi J.

; APPLICANT: Baker, Kevin P.

; APPLICANT: Botstein, David

; APPLICANT: Desnoyers, Luc

; APPLICANT: Eaton, Dan L.

; APPLICANT: Ferrara, Napoleone

; APPLICANT: Fong, Sherman

; APPLICANT: Gerber, Hanspeter

; APPLICANT: Gerritsen, Mary E.

; APPLICANT: Goddard, Audrey

; APPLICANT: Godowski, Paul J.

; APPLICANT: Grimaldi, J. Christopher

; APPLICANT: Gurney, Austin L.

; APPLICANT: Kljavin, Ivar J.

; APPLICANT: Napier, Mary A.

; APPLICANT: Pan, James

; APPLICANT: Paoni, Nicholas F.

; APPLICANT: Roy, Margaret Ann

; APPLICANT: Stewart, Timothy A.

; APPLICANT: Tumas, Daniel

; APPLICANT: Watanabe, Colin K.

; APPLICANT: Williams, P. Mickey

; APPLICANT: Wood, William I.

; APPLICANT: Zhang, Zemin

; TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic

; FILE REFERENCE: Acids Encoding the Same

; CURRENT APPLICATION NUMBER: US/09/992,598

; PRIOR FILING DATE: 2001-11-14

; PRIOR APPLICATION NUMBER: 60/049787

; PRIOR FILING DATE: 1997-06-16

[illegible]

; PRIOR APPLICATION NUMBER: 60/090863
; PRIOR FILING DATE: 1998-06-26
; PRIOR APPLICATION NUMBER: 60/091360
; PRIOR FILING DATE: 1998-07-01
; PRIOR APPLICATION NUMBER: 60/091478
; PRIOR FILING DATE: 1998-07-02
; PRIOR APPLICATION NUMBER: 60/091544
; PRIOR FILING DATE: 1998-07-01
; PRIOR APPLICATION NUMBER: 60/091519
; PRIOR FILING DATE: 1998-07-02
; PRIOR APPLICATION NUMBER: 60/091626
; PRIOR FILING DATE: 1998-07-02
; PRIOR APPLICATION NUMBER: 60/091633
; PRIOR FILING DATE: 1998-07-02
; PRIOR APPLICATION NUMBER: 60/091978
; PRIOR FILING DATE: 1998-07-07
; PRIOR APPLICATION NUMBER: 60/091982
; PRIOR FILING DATE: 1998-07-07
; PRIOR APPLICATION NUMBER: 60/092182
; PRIOR FILING DATE: 1998-07-09

Alignment Scores:

Pred. No.:
Score: 1.9e-05
Percent Similarity: 112.00
Best Local Similarity: 53.57%
Query Match: 42.86%
DB: 34.25%
Mismatches: 24
Indels: 2
Gaps: 2
Length: 762
Matches: 24
Conservative: 6

09-833799-13B (1-57) x US-09-992-598-344 (1-762)
QY 1 AlaGlnGluProValLysGlyProValSerThrLysProGlySerCysProIleIleLeu 20
Db 62 GCTGTGGAAGGAGTTAAAGAGGGTATAGAG---AAAGCAGGGGTTGCCAGCTGACAAC 118
QY 21 IleArgCysAlaMetLeuAsnProProAsnArgCysLeuLysAspThrAspCysProGly 40
Db 119 GTACGCTGCTTCAAGTCGATCCTCCC---CAAGTGCACACAGACCAGGACTGTCTGGG 175
QY 41 IleLysLysCysGluGlySerCysGlyMetAlaCysPheValPro 56
Db 176 GAAAGGAGGTGTTGTTACCTGCACCTGTGGCTTCAAGTGTGTGATTCCT 223

RESULT 9

US-09-989-293A-344
; Sequence 344, Application US/09989293A
; Patent No. US20020177164A1
; GENERAL INFORMATION:
; APPLICANT: Ashkenazi, Avi J.
; APPLICANT: Baker, Kevin P.
; APPLICANT: Botstein, David
; APPLICANT: Desnoyers, Luc
; APPLICANT: Eaton, Dan L.
; APPLICANT: Ferrara, Napoleone
; APPLICANT: Fong, Sherman
; APPLICANT: Gerber, Hanspeter
; APPLICANT: Gerritsen, Mary E.
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Grimaldi, J. Christopher
; APPLICANT: Gurney, Austin L.
; APPLICANT: Kljavin, Ivar J.
; APPLICANT: Napier, Mary A.
; APPLICANT: Pan, James
; APPLICANT: Paoni, Nicholas F.
; APPLICANT: Roy, Margaret Ann
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Watanabe, Colin K.
; APPLICANT: Williams, P. Mickey
; APPLICANT: Wood, William I.
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic

; TITLE OF INVENTION: Acids Encoding the Same
; FILE REFERENCE: P2730P1C66
; CURRENT APPLICATION NUMBER: US/09/989,293A
; CURRENT FILING DATE: 2001-11-20
; PRIOR APPLICATION NUMBER: 60/049787
; PRIOR FILING DATE: 1997-06-16
; PRIOR APPLICATION NUMBER: 60/062250
; PRIOR FILING DATE: 1997-10-17
; PRIOR APPLICATION NUMBER: 60/065186
; PRIOR FILING DATE: 1997-11-12
; PRIOR APPLICATION NUMBER: 60/065311
; PRIOR FILING DATE: 1997-11-13
; PRIOR APPLICATION NUMBER: 60/066770
; PRIOR FILING DATE: 1997-11-24
; PRIOR APPLICATION NUMBER: 60/075945
; PRIOR FILING DATE: 1998-02-25
; PRIOR APPLICATION NUMBER: 60/078910
; PRIOR FILING DATE: 1998-03-20
; PRIOR APPLICATION NUMBER: 60/083322
; PRIOR FILING DATE: 1998-04-28
; PRIOR APPLICATION NUMBER: 60/084600
; PRIOR FILING DATE: 1998-05-07
; PRIOR APPLICATION NUMBER: 60/087106
; PRIOR FILING DATE: 1998-05-28
; PRIOR APPLICATION NUMBER: 60/087607
; PRIOR FILING DATE: 1998-06-02
; PRIOR APPLICATION NUMBER: 60/087609
; PRIOR FILING DATE: 1998-06-02
; PRIOR APPLICATION NUMBER: 60/087759
; PRIOR FILING DATE: 1998-06-02
; PRIOR APPLICATION NUMBER: 60/087827
; PRIOR FILING DATE: 1998-06-03
; PRIOR APPLICATION NUMBER: 60/088021
; PRIOR FILING DATE: 1998-06-04
; PRIOR APPLICATION NUMBER: 60/088025
; PRIOR FILING DATE: 1998-06-04
; PRIOR APPLICATION NUMBER: 60/088026
; PRIOR FILING DATE: 1998-06-04
; PRIOR APPLICATION NUMBER: 60/088028
; PRIOR FILING DATE: 1998-06-04
; PRIOR APPLICATION NUMBER: 60/088029
; PRIOR FILING DATE: 1998-06-04
; PRIOR APPLICATION NUMBER: 60/088030
; PRIOR FILING DATE: 1998-06-04
; PRIOR APPLICATION NUMBER: 60/088033
; PRIOR FILING DATE: 1998-06-04
; PRIOR APPLICATION NUMBER: 60/088326
; PRIOR FILING DATE: 1998-06-04
; PRIOR APPLICATION NUMBER: 60/088167
; PRIOR FILING DATE: 1998-06-05
; PRIOR APPLICATION NUMBER: 60/088202
; PRIOR FILING DATE: 1998-06-05
; PRIOR APPLICATION NUMBER: 60/088212
; PRIOR FILING DATE: 1998-06-05
; PRIOR APPLICATION NUMBER: 60/088217
; PRIOR FILING DATE: 1998-06-05
; PRIOR APPLICATION NUMBER: 60/088655
; PRIOR FILING DATE: 1998-06-09
; PRIOR APPLICATION NUMBER: 60/088734
; PRIOR FILING DATE: 1998-06-10
; PRIOR APPLICATION NUMBER: 60/088738
; PRIOR FILING DATE: 1998-06-10
; PRIOR APPLICATION NUMBER: 60/088742
; PRIOR FILING DATE: 1998-06-10
; PRIOR APPLICATION NUMBER: 60/088810
; PRIOR FILING DATE: 1998-06-10
; PRIOR APPLICATION NUMBER: 60/088824
; PRIOR FILING DATE: 1998-06-10
; PRIOR APPLICATION NUMBER: 60/088826
; PRIOR FILING DATE: 1998-06-10
; PRIOR APPLICATION NUMBER: 60/088858
; PRIOR FILING DATE: 1998-06-11
; PRIOR APPLICATION NUMBER: 60/088861

09-833799-13b.rnpb

Sun Feb 16 09:13:26 2003

;; PRIOR APPLICATION NUMBER: 60/090695
;; PRIOR FILING DATE: 1998-06-25
;; PRIOR APPLICATION NUMBER: 60/090696
;; PRIOR FILING DATE: 1998-06-25
;; PRIOR APPLICATION NUMBER: 60/090862
;; PRIOR FILING DATE: 1998-06-26
;; PRIOR APPLICATION NUMBER: 60/090863
;; PRIOR FILING DATE: 1998-06-26
;; PRIOR APPLICATION NUMBER: 60/091360
;; PRIOR FILING DATE: 1998-07-01
;; PRIOR APPLICATION NUMBER: 60/091478
;; PRIOR FILING DATE: 1998-07-02
;; PRIOR APPLICATION NUMBER: 60/091544
;; PRIOR FILING DATE: 1998-07-01
;; PRIOR APPLICATION NUMBER: 60/091519
;; PRIOR FILING DATE: 1998-07-02
;; PRIOR APPLICATION NUMBER: 60/091626
;; PRIOR FILING DATE: 1998-07-02
;; PRIOR APPLICATION NUMBER: 60/091633
;; PRIOR FILING DATE: 1998-07-02
;; PRIOR APPLICATION NUMBER: 60/091978
;; PRIOR FILING DATE: 1998-07-07
;; PRIOR APPLICATION NUMBER: 60/091982
;; PRIOR FILING DATE: 1998-07-07
;; PRIOR APPLICATION NUMBER: 60/092182
;; PRIOR FILING DATE: 1998-07-09

Alignment Scores: 1.9e-05 762
Pred. No.: 112.00 24
Score: 53.57% 6
Percent Similarity: 42.86% 24
Best Local Similarity: 34.25% 2
Query Match: 9 2
DB: 2

09-833799-13B (1-57) x US-09-989-293A-344 (1-762)

QY 1 AlaGlnGluProValLysGlyProValSerThrLysProGlySerCysProIleLeu 20
Db 62 GCTGTGGAAGAGGTTAAAGAGGGTATAGAG---AAAGCAGGGGTTGCCAGCTGACAAC 118
QY 21 IleArgCysAlaMetLeuAsnProProAsnArgCysLeuLysAspThrAspCysProGly 40
Db 119 GTACGCTGCTTCAAGTCCGATCCTCCC---CAGTGTACACAGACCAGGACTGTCTGGGG 175
QY 41 IleLysLysCysCysGluGlySerCysGlyMetAlaCysPheValPro 56
Db 176 GAAAGGAAGTGTGTTTACCTGCTGCTGCTTCAAGTGTGTGATTCT 223

RESULT 10

US-09-989-735-344

; Sequence 344, Application US/09989735

; Publication No. US20020193299A1

; GENERAL INFORMATION:

; APPLICANT: Ashkenazi, Avi J.

; APPLICANT: Baker, Kevin P.

; APPLICANT: Botstein, David

; APPLICANT: Desnoyers, Luc

; APPLICANT: Eaton, Dan L.

; APPLICANT: Ferrara, Napoleone

; APPLICANT: Fong, Sherman

; APPLICANT: Gerber, Hanspeter

; APPLICANT: Gerritsen, Mary E.

; APPLICANT: Goddard, Audrey

; APPLICANT: Godowski, Paul J.

; APPLICANT: Grimaldi, J. Christopher

; APPLICANT: Gurney, Austin L.

; APPLICANT: Kljavin, Ivar J.

; APPLICANT: Napier, Mary A.

; APPLICANT: Pan, James

; APPLICANT: Paoni, Nicholas F.

; APPLICANT: Roy, Margaret Ann

; APPLICANT: Stewart, Timothy A.

;; PRIOR FILING DATE: 1998-06-11
;; PRIOR APPLICATION NUMBER: 60/088876
;; PRIOR FILING DATE: 1998-06-11
;; PRIOR APPLICATION NUMBER: 60/089105
;; PRIOR FILING DATE: 1998-06-12
;; PRIOR APPLICATION NUMBER: 60/089440
;; PRIOR FILING DATE: 1998-06-16
;; PRIOR APPLICATION NUMBER: 60/089512
;; PRIOR FILING DATE: 1998-06-16
;; PRIOR APPLICATION NUMBER: 60/089514
;; PRIOR FILING DATE: 1998-06-16
;; PRIOR APPLICATION NUMBER: 60/089532
;; PRIOR FILING DATE: 1998-06-17
;; PRIOR APPLICATION NUMBER: 60/089538
;; PRIOR FILING DATE: 1998-06-17
;; PRIOR APPLICATION NUMBER: 60/089598
;; PRIOR FILING DATE: 1998-06-17
;; PRIOR APPLICATION NUMBER: 60/089599
;; PRIOR FILING DATE: 1998-06-17
;; PRIOR APPLICATION NUMBER: 60/089600
;; PRIOR FILING DATE: 1998-06-17
;; PRIOR APPLICATION NUMBER: 60/089653
;; PRIOR FILING DATE: 1998-06-17
;; PRIOR APPLICATION NUMBER: 60/089801
;; PRIOR FILING DATE: 1998-06-18
;; PRIOR APPLICATION NUMBER: 60/089907
;; PRIOR FILING DATE: 1998-06-18
;; PRIOR APPLICATION NUMBER: 60/089908
;; PRIOR FILING DATE: 1998-06-18
;; PRIOR APPLICATION NUMBER: 60/089947
;; PRIOR FILING DATE: 1998-06-19
;; PRIOR APPLICATION NUMBER: 60/089948
;; PRIOR FILING DATE: 1998-06-19
;; PRIOR APPLICATION NUMBER: 60/089952
;; PRIOR FILING DATE: 1998-06-19
;; PRIOR APPLICATION NUMBER: 60/090246
;; PRIOR FILING DATE: 1998-06-22
;; PRIOR APPLICATION NUMBER: 60/090252
;; PRIOR FILING DATE: 1998-06-22
;; PRIOR APPLICATION NUMBER: 60/090254
;; PRIOR FILING DATE: 1998-06-22
;; PRIOR APPLICATION NUMBER: 60/090349
;; PRIOR FILING DATE: 1998-06-23
;; PRIOR APPLICATION NUMBER: 60/090355
;; PRIOR FILING DATE: 1998-06-23
;; PRIOR APPLICATION NUMBER: 60/090429
;; PRIOR FILING DATE: 1998-06-24
;; PRIOR APPLICATION NUMBER: 60/090431
;; PRIOR FILING DATE: 1998-06-24
;; PRIOR APPLICATION NUMBER: 60/090435
;; PRIOR FILING DATE: 1998-06-24
;; PRIOR APPLICATION NUMBER: 60/090444
;; PRIOR FILING DATE: 1998-06-24
;; PRIOR APPLICATION NUMBER: 60/090445
;; PRIOR FILING DATE: 1998-06-24
;; PRIOR APPLICATION NUMBER: 60/090472
;; PRIOR FILING DATE: 1998-06-24
;; PRIOR APPLICATION NUMBER: 60/090535
;; PRIOR FILING DATE: 1998-06-24
;; PRIOR APPLICATION NUMBER: 60/090540
;; PRIOR FILING DATE: 1998-06-24
;; PRIOR APPLICATION NUMBER: 60/090542
;; PRIOR FILING DATE: 1998-06-24
;; PRIOR APPLICATION NUMBER: 60/090557
;; PRIOR FILING DATE: 1998-06-24
;; PRIOR APPLICATION NUMBER: 60/090676
;; PRIOR FILING DATE: 1998-06-25
;; PRIOR APPLICATION NUMBER: 60/090678
;; PRIOR FILING DATE: 1998-06-25
;; PRIOR APPLICATION NUMBER: 60/090690
;; PRIOR FILING DATE: 1998-06-25
;; PRIOR APPLICATION NUMBER: 60/090694
;; PRIOR FILING DATE: 1998-06-25

APPLICANT: Tumas, Daniel
APPLICANT: Watanabe, Colin K.
APPLICANT: Williams, P. Mickey
APPLICANT: Wood, William I.
APPLICANT: Zhang, Zemin
TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
FILE REFERENCE: P2730P1C61
CURRENT APPLICATION NUMBER: US/09/989,735
PRIOR FILING DATE: 2001-11-19
PRIOR APPLICATION NUMBER: 60/049787
PRIOR FILING DATE: 1997-06-16
PRIOR APPLICATION NUMBER: 60/062250
PRIOR FILING DATE: 1997-10-17
PRIOR APPLICATION NUMBER: 60/065186
PRIOR FILING DATE: 1997-11-12
PRIOR APPLICATION NUMBER: 60/065311
PRIOR FILING DATE: 1997-11-13
PRIOR APPLICATION NUMBER: 60/066770
PRIOR FILING DATE: 1997-11-24
PRIOR APPLICATION NUMBER: 60/075945
PRIOR FILING DATE: 1998-02-25
PRIOR APPLICATION NUMBER: 60/078910
PRIOR FILING DATE: 1998-03-20
PRIOR APPLICATION NUMBER: 60/083322
PRIOR FILING DATE: 1998-04-28
PRIOR APPLICATION NUMBER: 60/084600
PRIOR FILING DATE: 1998-05-07
PRIOR APPLICATION NUMBER: 60/087106
PRIOR FILING DATE: 1998-05-28
PRIOR APPLICATION NUMBER: 60/087607
PRIOR FILING DATE: 1998-06-02
PRIOR APPLICATION NUMBER: 60/087609
PRIOR FILING DATE: 1998-06-02
PRIOR APPLICATION NUMBER: 60/087759
PRIOR FILING DATE: 1998-06-02
PRIOR APPLICATION NUMBER: 60/087827
PRIOR FILING DATE: 1998-06-03
PRIOR APPLICATION NUMBER: 60/088021
PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088025
PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088026
PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088028
PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088029
PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088030
PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088033
PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088326
PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088167
PRIOR FILING DATE: 1998-06-05
PRIOR APPLICATION NUMBER: 60/088202
PRIOR FILING DATE: 1998-06-05
PRIOR APPLICATION NUMBER: 60/088212
PRIOR FILING DATE: 1998-06-05
PRIOR APPLICATION NUMBER: 60/088217
PRIOR FILING DATE: 1998-06-05
PRIOR APPLICATION NUMBER: 60/088655
PRIOR FILING DATE: 1998-06-09
PRIOR APPLICATION NUMBER: 60/088734
PRIOR FILING DATE: 1998-06-10
PRIOR APPLICATION NUMBER: 60/088738
PRIOR FILING DATE: 1998-06-10
PRIOR APPLICATION NUMBER: 60/088742
PRIOR FILING DATE: 1998-06-10
PRIOR APPLICATION NUMBER: 60/088810
PRIOR FILING DATE: 1998-06-10
PRIOR APPLICATION NUMBER: 60/088824

PRIOR FILING DATE: 1998-06-10
PRIOR APPLICATION NUMBER: 60/088826
PRIOR FILING DATE: 1998-06-10
PRIOR APPLICATION NUMBER: 60/088858
PRIOR FILING DATE: 1998-06-11
PRIOR APPLICATION NUMBER: 60/088861
PRIOR FILING DATE: 1998-06-11
PRIOR APPLICATION NUMBER: 60/088876
PRIOR FILING DATE: 1998-06-11
PRIOR APPLICATION NUMBER: 60/089105
PRIOR FILING DATE: 1998-06-12
PRIOR APPLICATION NUMBER: 60/089440
PRIOR FILING DATE: 1998-06-16
PRIOR APPLICATION NUMBER: 60/089512
PRIOR FILING DATE: 1998-06-16
PRIOR APPLICATION NUMBER: 60/089514
PRIOR FILING DATE: 1998-06-16
PRIOR APPLICATION NUMBER: 60/089532
PRIOR FILING DATE: 1998-06-17
PRIOR APPLICATION NUMBER: 60/089538
PRIOR FILING DATE: 1998-06-17
PRIOR APPLICATION NUMBER: 60/089598
PRIOR FILING DATE: 1998-06-17
PRIOR APPLICATION NUMBER: 60/089599
PRIOR FILING DATE: 1998-06-17
PRIOR APPLICATION NUMBER: 60/089600
PRIOR FILING DATE: 1998-06-17
PRIOR APPLICATION NUMBER: 60/089653
PRIOR FILING DATE: 1998-06-17
PRIOR APPLICATION NUMBER: 60/089801
PRIOR FILING DATE: 1998-06-18
PRIOR APPLICATION NUMBER: 60/089907
PRIOR FILING DATE: 1998-06-18
PRIOR APPLICATION NUMBER: 60/089908
PRIOR FILING DATE: 1998-06-18
PRIOR APPLICATION NUMBER: 60/089947
PRIOR FILING DATE: 1998-06-19
PRIOR APPLICATION NUMBER: 60/089948
PRIOR FILING DATE: 1998-06-19
PRIOR APPLICATION NUMBER: 60/089952
PRIOR FILING DATE: 1998-06-19
PRIOR APPLICATION NUMBER: 60/090246
PRIOR FILING DATE: 1998-06-22
PRIOR APPLICATION NUMBER: 60/090252
PRIOR FILING DATE: 1998-06-22
PRIOR APPLICATION NUMBER: 60/090254
PRIOR FILING DATE: 1998-06-22
PRIOR APPLICATION NUMBER: 60/090349
PRIOR FILING DATE: 1998-06-23
PRIOR APPLICATION NUMBER: 60/090355
PRIOR FILING DATE: 1998-06-23
PRIOR APPLICATION NUMBER: 60/090429
PRIOR FILING DATE: 1998-06-24
PRIOR APPLICATION NUMBER: 60/090431
PRIOR FILING DATE: 1998-06-24
PRIOR APPLICATION NUMBER: 60/090435
PRIOR FILING DATE: 1998-06-24
PRIOR APPLICATION NUMBER: 60/090444
PRIOR FILING DATE: 1998-06-24
PRIOR APPLICATION NUMBER: 60/090445
PRIOR FILING DATE: 1998-06-24
PRIOR APPLICATION NUMBER: 60/090472
PRIOR FILING DATE: 1998-06-24
PRIOR APPLICATION NUMBER: 60/090535
PRIOR FILING DATE: 1998-06-24
PRIOR APPLICATION NUMBER: 60/090540
PRIOR FILING DATE: 1998-06-24
PRIOR APPLICATION NUMBER: 60/090542
PRIOR FILING DATE: 1998-06-24
PRIOR APPLICATION NUMBER: 60/090557
PRIOR FILING DATE: 1998-06-24
PRIOR APPLICATION NUMBER: 60/090676
PRIOR FILING DATE: 1998-06-25

; PRIOR APPLICATION NUMBER: 60/090678
; PRIOR FILING DATE: 1998-06-25
; PRIOR APPLICATION NUMBER: 60/090690
; PRIOR FILING DATE: 1998-06-25
; PRIOR APPLICATION NUMBER: 60/090694
; PRIOR FILING DATE: 1998-06-25
; PRIOR APPLICATION NUMBER: 60/090695
; PRIOR FILING DATE: 1998-06-25
; PRIOR APPLICATION NUMBER: 60/090696
; PRIOR FILING DATE: 1998-06-25
; PRIOR APPLICATION NUMBER: 60/090862
; PRIOR FILING DATE: 1998-06-26
; PRIOR APPLICATION NUMBER: 60/090863
; PRIOR FILING DATE: 1998-06-26
; PRIOR APPLICATION NUMBER: 60/091360
; PRIOR FILING DATE: 1998-07-01
; PRIOR APPLICATION NUMBER: 60/091478
; PRIOR FILING DATE: 1998-07-02
; PRIOR APPLICATION NUMBER: 60/091544
; PRIOR FILING DATE: 1998-07-01
; PRIOR APPLICATION NUMBER: 60/091519
; PRIOR FILING DATE: 1998-07-02
; PRIOR APPLICATION NUMBER: 60/091626
; PRIOR FILING DATE: 1998-07-02
; PRIOR APPLICATION NUMBER: 60/091633
; PRIOR FILING DATE: 1998-07-02
; PRIOR APPLICATION NUMBER: 60/091978
; PRIOR FILING DATE: 1998-07-07
; PRIOR APPLICATION NUMBER: 60/091982
; PRIOR FILING DATE: 1998-07-07
; PRIOR APPLICATION NUMBER: 60/092182
; PRIOR FILING DATE: 1998-07-09

Alignment Scores:

Pred. No.: 1.9e-05 Length: 762
Score: 112.00 Matches: 24
Percent Similarity: 53.57% Conservative: 6
Best Local Similarity: 42.86% Mismatches: 24
Query Match: 34.25% Indels: 2
DB: 9 Gaps: 2

09-833799-13B (1-57) x US-09-989-735-344 (1-762)

QY 1 AlaGlnGluProValLysGlyProValSerThrLysProGlySerCysProIleLeu 20

Db 62 GCTGTGAAGGAGTTAAAGAGGGTATAGAG--AAAGCAGGGGTTTGCACAGTGCACAAC 118

QY 21 IleArgCysAlaMetLeuAsnProProAsnArgCysLeuLysAspThrAspCysProGly 40

Db 119 GTACGCTGCTTCAAGTCCGATCCTCCC---CAGTGTACACAGACCAGGACTGTCTGGGG 175

QY 41 IleLysLysCysCysGluGlySerCysGlyMetAlaCysPheValPro 56

Db 176 GAAAGGAAGTGTGTACCTGCACTGTGGCTTCAAGTGTGTGATTCTCT 223

RESULT 11

US-09-990-444-344

; Sequence 344, Application US/09990444

; Publication No. US20020193300A1

GENERAL INFORMATION:

; APPLICANT: Ashkenazi, Avi J.
; APPLICANT: Baker, Kevin P.
; APPLICANT: Botstein, David
; APPLICANT: Desnoyers, Luc
; APPLICANT: Eaton, Dan L.
; APPLICANT: Ferrara, Napoleone
; APPLICANT: Fong, Sherman
; APPLICANT: Gerber, Hanspeter
; APPLICANT: Gerritsen, Mary E.
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Grimaldi, J. Christopher
; APPLICANT: Gurney, Austin L.

; APPLICANT: Kljavin, Ivar J.
; APPLICANT: Napier, Mary A.
; APPLICANT: Pan, James
; APPLICANT: Paoni, Nicholas F.
; APPLICANT: Roy, Margaret Ann
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Watanabe, Colin K.
; APPLICANT: Williams, P. Mickey
; APPLICANT: Wood, William I.
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
; FILE REFERENCE: P2730P1C19
; CURRENT APPLICATION NUMBER: US/09/990,444
; PRIOR FILING DATE: 2001-11-14
; PRIOR APPLICATION NUMBER: 60/049787
; PRIOR FILING DATE: 1997-06-16
; PRIOR APPLICATION NUMBER: 60/062250
; PRIOR FILING DATE: 1997-10-17
; PRIOR APPLICATION NUMBER: 60/065186
; PRIOR FILING DATE: 1997-11-12
; PRIOR APPLICATION NUMBER: 60/065311
; PRIOR FILING DATE: 1997-11-13
; PRIOR APPLICATION NUMBER: 60/066770
; PRIOR FILING DATE: 1997-11-24
; PRIOR APPLICATION NUMBER: 60/075945
; PRIOR FILING DATE: 1998-02-25
; PRIOR APPLICATION NUMBER: 60/078910
; PRIOR FILING DATE: 1998-03-20
; PRIOR APPLICATION NUMBER: 60/083322
; PRIOR FILING DATE: 1998-04-28
; PRIOR APPLICATION NUMBER: 60/084600
; PRIOR FILING DATE: 1998-05-07
; PRIOR APPLICATION NUMBER: 60/087106
; PRIOR FILING DATE: 1998-05-28
; PRIOR APPLICATION NUMBER: 60/087607
; PRIOR FILING DATE: 1998-06-02
; PRIOR APPLICATION NUMBER: 60/087609
; PRIOR FILING DATE: 1998-06-02
; PRIOR APPLICATION NUMBER: 60/087759
; PRIOR FILING DATE: 1998-06-02
; PRIOR APPLICATION NUMBER: 60/087827
; PRIOR FILING DATE: 1998-06-03
; PRIOR APPLICATION NUMBER: 60/088021
; PRIOR FILING DATE: 1998-06-04
; PRIOR APPLICATION NUMBER: 60/088025
; PRIOR FILING DATE: 1998-06-04
; PRIOR APPLICATION NUMBER: 60/088026
; PRIOR FILING DATE: 1998-06-04
; PRIOR APPLICATION NUMBER: 60/088028
; PRIOR FILING DATE: 1998-06-04
; PRIOR APPLICATION NUMBER: 60/088029
; PRIOR FILING DATE: 1998-06-04
; PRIOR APPLICATION NUMBER: 60/088030
; PRIOR FILING DATE: 1998-06-04
; PRIOR APPLICATION NUMBER: 60/088033
; PRIOR FILING DATE: 1998-06-04
; PRIOR APPLICATION NUMBER: 60/088326
; PRIOR FILING DATE: 1998-06-04
; PRIOR APPLICATION NUMBER: 60/088167
; PRIOR FILING DATE: 1998-06-05
; PRIOR APPLICATION NUMBER: 60/088202
; PRIOR FILING DATE: 1998-06-05
; PRIOR APPLICATION NUMBER: 60/088212
; PRIOR FILING DATE: 1998-06-05
; PRIOR APPLICATION NUMBER: 60/088217
; PRIOR FILING DATE: 1998-06-05
; PRIOR APPLICATION NUMBER: 60/088655
; PRIOR FILING DATE: 1998-06-09
; PRIOR APPLICATION NUMBER: 60/088734
; PRIOR FILING DATE: 1998-06-10
; PRIOR APPLICATION NUMBER: 60/088738

; PRIOR FILING DATE: 1998-06-10
; PRIOR APPLICATION NUMBER: 60/088742
; PRIOR FILING DATE: 1998-06-10
; PRIOR APPLICATION NUMBER: 60/088810
; PRIOR FILING DATE: 1998-06-10
; PRIOR APPLICATION NUMBER: 60/088824
; PRIOR FILING DATE: 1998-06-10
; PRIOR APPLICATION NUMBER: 60/088826
; PRIOR FILING DATE: 1998-06-10
; PRIOR APPLICATION NUMBER: 60/088858
; PRIOR FILING DATE: 1998-06-11
; PRIOR APPLICATION NUMBER: 60/088861
; PRIOR FILING DATE: 1998-06-11
; PRIOR APPLICATION NUMBER: 60/088876
; PRIOR FILING DATE: 1998-06-11
; PRIOR APPLICATION NUMBER: 60/089105
; PRIOR FILING DATE: 1998-06-12
; PRIOR APPLICATION NUMBER: 60/089440
; PRIOR FILING DATE: 1998-06-16
; PRIOR APPLICATION NUMBER: 60/089512
; PRIOR FILING DATE: 1998-06-16
; PRIOR APPLICATION NUMBER: 60/089514
; PRIOR FILING DATE: 1998-06-16
; PRIOR APPLICATION NUMBER: 60/089532
; PRIOR FILING DATE: 1998-06-17
; PRIOR APPLICATION NUMBER: 60/089538
; PRIOR FILING DATE: 1998-06-17
; PRIOR APPLICATION NUMBER: 60/089598
; PRIOR FILING DATE: 1998-06-17
; PRIOR APPLICATION NUMBER: 60/089599
; PRIOR FILING DATE: 1998-06-17
; PRIOR APPLICATION NUMBER: 60/089600
; PRIOR FILING DATE: 1998-06-17
; PRIOR APPLICATION NUMBER: 60/089653
; PRIOR FILING DATE: 1998-06-17
; PRIOR APPLICATION NUMBER: 60/089801
; PRIOR FILING DATE: 1998-06-18
; PRIOR APPLICATION NUMBER: 60/089907
; PRIOR FILING DATE: 1998-06-18
; PRIOR APPLICATION NUMBER: 60/089908
; PRIOR FILING DATE: 1998-06-18
; PRIOR APPLICATION NUMBER: 60/089947
; PRIOR FILING DATE: 1998-06-19
; PRIOR APPLICATION NUMBER: 60/089948
; PRIOR FILING DATE: 1998-06-19
; PRIOR APPLICATION NUMBER: 60/089952
; PRIOR FILING DATE: 1998-06-19
; PRIOR APPLICATION NUMBER: 60/090246
; PRIOR FILING DATE: 1998-06-22
; PRIOR APPLICATION NUMBER: 60/090252
; PRIOR FILING DATE: 1998-06-22
; PRIOR APPLICATION NUMBER: 60/090254
; PRIOR FILING DATE: 1998-06-22
; PRIOR APPLICATION NUMBER: 60/090349
; PRIOR FILING DATE: 1998-06-23
; PRIOR APPLICATION NUMBER: 60/090355
; PRIOR FILING DATE: 1998-06-23
; PRIOR APPLICATION NUMBER: 60/090429
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090431
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090435
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090444
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090445
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090472
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090535
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090540
; PRIOR FILING DATE: 1998-06-24

; PRIOR APPLICATION NUMBER: 60/090542
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090557
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090676
; PRIOR FILING DATE: 1998-06-25
; PRIOR APPLICATION NUMBER: 60/090678
; PRIOR FILING DATE: 1998-06-25
; PRIOR APPLICATION NUMBER: 60/090690
; PRIOR FILING DATE: 1998-06-25
; PRIOR APPLICATION NUMBER: 60/090694
; PRIOR FILING DATE: 1998-06-25
; PRIOR APPLICATION NUMBER: 60/090695
; PRIOR FILING DATE: 1998-06-25
; PRIOR APPLICATION NUMBER: 60/090696
; PRIOR FILING DATE: 1998-06-25
; PRIOR APPLICATION NUMBER: 60/090862
; PRIOR FILING DATE: 1998-06-26
; PRIOR APPLICATION NUMBER: 60/090863
; PRIOR FILING DATE: 1998-06-26
; PRIOR APPLICATION NUMBER: 60/091360
; PRIOR FILING DATE: 1998-07-01
; PRIOR APPLICATION NUMBER: 60/091478
; PRIOR FILING DATE: 1998-07-02
; PRIOR APPLICATION NUMBER: 60/091544
; PRIOR FILING DATE: 1998-07-01
; PRIOR APPLICATION NUMBER: 60/091519
; PRIOR FILING DATE: 1998-07-02
; PRIOR APPLICATION NUMBER: 60/091626
; PRIOR FILING DATE: 1998-07-02
; PRIOR APPLICATION NUMBER: 60/091633
; PRIOR FILING DATE: 1998-07-02
; PRIOR APPLICATION NUMBER: 60/091978
; PRIOR FILING DATE: 1998-07-07
; PRIOR APPLICATION NUMBER: 60/091982
; PRIOR FILING DATE: 1998-07-07
; PRIOR APPLICATION NUMBER: 60/092182
; PRIOR FILING DATE: 1998-07-09

Alignment Scores:

Pred. No.:	1.9e-05	Length:	762
Score:	112.00	Matches:	24
Percent Similarity:	53.57%	Conservative:	6
Best Local Similarity:	42.86%	Mismatches:	24
Query Match:	34.25%	Indels:	2
DB:	9	Gaps:	2

09-833799-13B (1-57) x US-09-990-444-344 (1-762)

Qy	1	AlaGlnGluProValLysGlyProValSerThrLysProGlySerCysProIleLeu	20
Db	62	GCTGTGAAGGAGTTAAAGAGGGTATAGAG--AAAGCAGGGGTTGCCAGCTGACAAC	118
Qy	21	IleArgCysAlaMetLeuAsnProProAsnArgCysLeuLysAspThrAspCysProGly	40
Db	119	GTACGCTGCTTCAAGTCCGATCCTCCC---CAGTGTACACACAGACCAGGACTGTCTGGGG	175
Qy	41	IleLysLysCysCysGluGlySerCysGlyMetAlaCysPheValPro	56
Db	176	GAAGGAAGTGTGTTACCTGCACTGTGGCTTCAAGTGTGTGATTCTCT	223

RESULT 12

US-09-989-730-344
; Sequence 344, Application US/09989730
; Publication No. US20020197674A1
; GENERAL INFORMATION:
; APPLICANT: Ashkenazi, Avi J.
; APPLICANT: Baker, Kevin P.
; APPLICANT: Botstein, David
; APPLICANT: Desnoyers, Luc
; APPLICANT: Eaton, Dan L.
; APPLICANT: Ferrara, Napoleone
; APPLICANT: Fong, Sherman

; APPLICANT: Gerber,Hanspeter
; APPLICANT: Gerritsen,Mary E.
; APPLICANT: Goddard,Audrey
; APPLICANT: Godowski,Paul J.
; APPLICANT: Grimaldi,J.Christopher
; APPLICANT: Gurney,Austin L.
; APPLICANT: Kljavin,Ivar J.
; APPLICANT: Napier,Mary A.
; APPLICANT: Pan,James
; APPLICANT: Paoni,Nicholas F.
; APPLICANT: Roy, Margaret Ann
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Watanabe, Colin K.
; APPLICANT: Williams, P. Mickey
; APPLICANT: Wood, William I.
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
; TITLE OF INVENTION: Acids Encoding the Same
; FILE REFERENCE: P2730P1C69
; CURRENT APPLICATION NUMBER: US/09/989,730
; CURRENT FILING DATE: 2001-11-20
; PRIOR APPLICATION NUMBER: 60/049787
; PRIOR FILING DATE: 1997-06-16
; PRIOR APPLICATION NUMBER: 60/062250
; PRIOR FILING DATE: 1997-10-17
; PRIOR APPLICATION NUMBER: 60/065186
; PRIOR FILING DATE: 1997-11-12
; PRIOR APPLICATION NUMBER: 60/065311
; PRIOR FILING DATE: 1997-11-13
; PRIOR APPLICATION NUMBER: 60/066770
; PRIOR FILING DATE: 1997-11-24
; PRIOR APPLICATION NUMBER: 60/075945
; PRIOR FILING DATE: 1998-02-25
; PRIOR APPLICATION NUMBER: 60/078910
; PRIOR FILING DATE: 1998-03-20
; PRIOR APPLICATION NUMBER: 60/083322
; PRIOR FILING DATE: 1998-04-28
; PRIOR APPLICATION NUMBER: 60/084600
; PRIOR FILING DATE: 1998-05-07
; PRIOR APPLICATION NUMBER: 60/087106
; PRIOR FILING DATE: 1998-05-28
; PRIOR APPLICATION NUMBER: 60/087607
; PRIOR FILING DATE: 1998-06-02
; PRIOR APPLICATION NUMBER: 60/087609
; PRIOR FILING DATE: 1998-06-02
; PRIOR APPLICATION NUMBER: 60/087759
; PRIOR FILING DATE: 1998-06-02
; PRIOR APPLICATION NUMBER: 60/087827
; PRIOR FILING DATE: 1998-06-03
; PRIOR APPLICATION NUMBER: 60/088021
; PRIOR FILING DATE: 1998-06-04
; PRIOR APPLICATION NUMBER: 60/088025
; PRIOR FILING DATE: 1998-06-04
; PRIOR APPLICATION NUMBER: 60/088026
; PRIOR FILING DATE: 1998-06-04
; PRIOR APPLICATION NUMBER: 60/088028
; PRIOR FILING DATE: 1998-06-04
; PRIOR APPLICATION NUMBER: 60/088029
; PRIOR FILING DATE: 1998-06-04
; PRIOR APPLICATION NUMBER: 60/088030
; PRIOR FILING DATE: 1998-06-04
; PRIOR APPLICATION NUMBER: 60/088033
; PRIOR FILING DATE: 1998-06-04
; PRIOR APPLICATION NUMBER: 60/088326
; PRIOR FILING DATE: 1998-06-04
; PRIOR APPLICATION NUMBER: 60/088167
; PRIOR FILING DATE: 1998-06-05
; PRIOR APPLICATION NUMBER: 60/088202
; PRIOR FILING DATE: 1998-06-05
; PRIOR APPLICATION NUMBER: 60/088212
; PRIOR FILING DATE: 1998-06-05
; PRIOR APPLICATION NUMBER: 60/088217
; PRIOR FILING DATE: 1998-06-05
; PRIOR APPLICATION NUMBER: 60/088655
; PRIOR FILING DATE: 1998-06-09
; PRIOR APPLICATION NUMBER: 60/088734
; PRIOR FILING DATE: 1998-06-10
; PRIOR APPLICATION NUMBER: 60/088738
; PRIOR FILING DATE: 1998-06-10
; PRIOR APPLICATION NUMBER: 60/088742
; PRIOR FILING DATE: 1998-06-10
; PRIOR APPLICATION NUMBER: 60/088810
; PRIOR FILING DATE: 1998-06-10
; PRIOR APPLICATION NUMBER: 60/088824
; PRIOR FILING DATE: 1998-06-10
; PRIOR APPLICATION NUMBER: 60/088826
; PRIOR FILING DATE: 1998-06-10
; PRIOR APPLICATION NUMBER: 60/088858
; PRIOR FILING DATE: 1998-06-11
; PRIOR APPLICATION NUMBER: 60/088861
; PRIOR FILING DATE: 1998-06-11
; PRIOR APPLICATION NUMBER: 60/088876
; PRIOR FILING DATE: 1998-06-11
; PRIOR APPLICATION NUMBER: 60/089105
; PRIOR FILING DATE: 1998-06-12
; PRIOR APPLICATION NUMBER: 60/089440
; PRIOR FILING DATE: 1998-06-16
; PRIOR APPLICATION NUMBER: 60/089512
; PRIOR FILING DATE: 1998-06-16
; PRIOR APPLICATION NUMBER: 60/089514
; PRIOR FILING DATE: 1998-06-16
; PRIOR APPLICATION NUMBER: 60/089532
; PRIOR FILING DATE: 1998-06-17
; PRIOR APPLICATION NUMBER: 60/089538
; PRIOR FILING DATE: 1998-06-17
; PRIOR APPLICATION NUMBER: 60/089598
; PRIOR FILING DATE: 1998-06-17
; PRIOR APPLICATION NUMBER: 60/089599
; PRIOR FILING DATE: 1998-06-17
; PRIOR APPLICATION NUMBER: 60/089600
; PRIOR FILING DATE: 1998-06-17
; PRIOR APPLICATION NUMBER: 60/089653
; PRIOR FILING DATE: 1998-06-17
; PRIOR APPLICATION NUMBER: 60/089801
; PRIOR FILING DATE: 1998-06-18
; PRIOR APPLICATION NUMBER: 60/089907
; PRIOR FILING DATE: 1998-06-18
; PRIOR APPLICATION NUMBER: 60/089908
; PRIOR FILING DATE: 1998-06-18
; PRIOR APPLICATION NUMBER: 60/089947
; PRIOR FILING DATE: 1998-06-19
; PRIOR APPLICATION NUMBER: 60/089948
; PRIOR FILING DATE: 1998-06-19
; PRIOR APPLICATION NUMBER: 60/089952
; PRIOR FILING DATE: 1998-06-19
; PRIOR APPLICATION NUMBER: 60/090246
; PRIOR FILING DATE: 1998-06-22
; PRIOR APPLICATION NUMBER: 60/090252
; PRIOR FILING DATE: 1998-06-22
; PRIOR APPLICATION NUMBER: 60/090254
; PRIOR FILING DATE: 1998-06-22
; PRIOR APPLICATION NUMBER: 60/090349
; PRIOR FILING DATE: 1998-06-23
; PRIOR APPLICATION NUMBER: 60/090355
; PRIOR FILING DATE: 1998-06-23
; PRIOR APPLICATION NUMBER: 60/090429
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090431
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090435
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090444
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090445
; PRIOR FILING DATE: 1998-06-24

; PRIOR APPLICATION NUMBER: 60/090472
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090535
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090540
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090542
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090557
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090676
; PRIOR FILING DATE: 1998-06-25
; PRIOR APPLICATION NUMBER: 60/090678
; PRIOR FILING DATE: 1998-06-25
; PRIOR APPLICATION NUMBER: 60/090690
; PRIOR FILING DATE: 1998-06-25
; PRIOR APPLICATION NUMBER: 60/090694
; PRIOR FILING DATE: 1998-06-25
; PRIOR APPLICATION NUMBER: 60/090695
; PRIOR FILING DATE: 1998-06-25
; PRIOR APPLICATION NUMBER: 60/090696
; PRIOR FILING DATE: 1998-06-25
; PRIOR APPLICATION NUMBER: 60/090862
; PRIOR FILING DATE: 1998-06-26
; PRIOR APPLICATION NUMBER: 60/091360
; PRIOR FILING DATE: 1998-07-01
; PRIOR APPLICATION NUMBER: 60/091478
; PRIOR FILING DATE: 1998-07-02
; PRIOR APPLICATION NUMBER: 60/091544
; PRIOR FILING DATE: 1998-07-01
; PRIOR APPLICATION NUMBER: 60/091519
; PRIOR FILING DATE: 1998-07-02
; PRIOR APPLICATION NUMBER: 60/091626
; PRIOR FILING DATE: 1998-07-02
; PRIOR APPLICATION NUMBER: 60/091633
; PRIOR FILING DATE: 1998-07-02
; PRIOR APPLICATION NUMBER: 60/091978
; PRIOR FILING DATE: 1998-07-07
; PRIOR APPLICATION NUMBER: 60/091982
; PRIOR FILING DATE: 1998-07-07
; PRIOR APPLICATION NUMBER: 60/092182
; PRIOR FILING DATE: 1998-07-09

Alignment Scores:
Pred. No.: 1.9e-05 Length: 762
Score: 112.00 Matches: 24
Percent Similarity: 53.57% Conservative: 6
Best Local Similarity: 42.86% Mismatches: 24
Query Match: 34.25% Indels: 2
DB: 9 Gaps: 2

09-833799-13B (1-57) x US-09-989-730-344 (1-762)
Qy 1 AlaGlnGluProValLysGlyProValSerThrLysProGlySerCysProIleLeu 20
Db 62 GCTGTGGAAGGAGTTAAAGAGGGGTATAGAG---AAAGCAGGGGTTTGCCCGAGCTGACAAC 118
Qy 21 IleArgCysAlaMetLeuAsnProProAsnArgCysLeuLysAspThrAspCysProGly 40
Db 119 GTACGCTGCTTCAAGTCCGATCCTCCC---CAGTGTACACAGACCAGGACTGTCTGGGG 175
Qy 41 IleLysLysCysCysGluGlySerCysGlyMetAlaCysPheValPro 56
Db 176 GAAAGGAAGTGTGTTTACCTGCACCTGTGGCTTCAAGTGTGTGATTCCT 223

RESULT 13
US-09-990-436-344
; Sequence 344, Application US/09990436
; Publication No. US20020198148A1
; GENERAL INFORMATION:
; APPLICANT: Ashkenazi, Avi J.

; APPLICANT: Baker, Kevin P.
; APPLICANT: Botstein, David
; APPLICANT: Desnoyers, Luc
; APPLICANT: Eaton, Dan L.
; APPLICANT: Ferrara, Napoleone
; APPLICANT: Fong, Sherman
; APPLICANT: Gerber, Hanspeter
; APPLICANT: Gerritsen, Mary E.
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Grimaldi, J. Christopher
; APPLICANT: Gurney, Austin L.
; APPLICANT: Kljavin, Ivar J.
; APPLICANT: Napier, Mary A.
; APPLICANT: Pan, James
; APPLICANT: Paoni, Nicholas F.
; APPLICANT: Roy, Margaret Ann
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Watanabe, Colin K.
; APPLICANT: Williams, P. Mickey
; APPLICANT: Wood, William I.
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
; FILE REFERENCE: P2730P1C14
; CURRENT APPLICATION NUMBER: US/09/990,436
; CURRENT FILING DATE: 2001-11-14
; PRIOR APPLICATION NUMBER: 60/049787
; PRIOR FILING DATE: 1997-06-16
; PRIOR APPLICATION NUMBER: 60/062250
; PRIOR FILING DATE: 1997-10-17
; PRIOR APPLICATION NUMBER: 60/065186
; PRIOR FILING DATE: 1997-11-12
; PRIOR APPLICATION NUMBER: 60/065311
; PRIOR FILING DATE: 1997-11-13
; PRIOR APPLICATION NUMBER: 60/066770
; PRIOR FILING DATE: 1997-11-24
; PRIOR APPLICATION NUMBER: 60/075945
; PRIOR FILING DATE: 1998-02-25
; PRIOR APPLICATION NUMBER: 60/078910
; PRIOR FILING DATE: 1998-03-20
; PRIOR APPLICATION NUMBER: 60/083322
; PRIOR FILING DATE: 1998-04-28
; PRIOR APPLICATION NUMBER: 60/084600
; PRIOR FILING DATE: 1998-05-07
; PRIOR APPLICATION NUMBER: 60/087106
; PRIOR FILING DATE: 1998-05-28
; PRIOR APPLICATION NUMBER: 60/087607
; PRIOR FILING DATE: 1998-06-02
; PRIOR APPLICATION NUMBER: 60/087609
; PRIOR FILING DATE: 1998-06-02
; PRIOR APPLICATION NUMBER: 60/087759
; PRIOR FILING DATE: 1998-06-02
; PRIOR APPLICATION NUMBER: 60/087827
; PRIOR FILING DATE: 1998-06-03
; PRIOR APPLICATION NUMBER: 60/088021
; PRIOR FILING DATE: 1998-06-04
; PRIOR APPLICATION NUMBER: 60/088025
; PRIOR FILING DATE: 1998-06-04
; PRIOR APPLICATION NUMBER: 60/088026
; PRIOR FILING DATE: 1998-06-04
; PRIOR APPLICATION NUMBER: 60/088028
; PRIOR FILING DATE: 1998-06-04
; PRIOR APPLICATION NUMBER: 60/088029
; PRIOR FILING DATE: 1998-06-04
; PRIOR APPLICATION NUMBER: 60/088030
; PRIOR FILING DATE: 1998-06-04
; PRIOR APPLICATION NUMBER: 60/088033
; PRIOR FILING DATE: 1998-06-04
; PRIOR APPLICATION NUMBER: 60/088326
; PRIOR FILING DATE: 1998-06-04
; PRIOR APPLICATION NUMBER: 60/088167

; PRIOR FILING DATE: 1998-06-05
; PRIOR APPLICATION NUMBER: 60/088202
; PRIOR FILING DATE: 1998-06-05
; PRIOR APPLICATION NUMBER: 60/088212
; PRIOR FILING DATE: 1998-06-05
; PRIOR APPLICATION NUMBER: 60/088217
; PRIOR FILING DATE: 1998-06-05
; PRIOR APPLICATION NUMBER: 60/088655
; PRIOR FILING DATE: 1998-06-09
; PRIOR APPLICATION NUMBER: 60/088734
; PRIOR FILING DATE: 1998-06-10
; PRIOR APPLICATION NUMBER: 60/088738
; PRIOR FILING DATE: 1998-06-10
; PRIOR APPLICATION NUMBER: 60/088742
; PRIOR FILING DATE: 1998-06-10
; PRIOR APPLICATION NUMBER: 60/088810
; PRIOR FILING DATE: 1998-06-10
; PRIOR APPLICATION NUMBER: 60/088824
; PRIOR FILING DATE: 1998-06-10
; PRIOR APPLICATION NUMBER: 60/088826
; PRIOR FILING DATE: 1998-06-10
; PRIOR APPLICATION NUMBER: 60/088858
; PRIOR FILING DATE: 1998-06-11
; PRIOR APPLICATION NUMBER: 60/088861
; PRIOR FILING DATE: 1998-06-11
; PRIOR APPLICATION NUMBER: 60/088876
; PRIOR FILING DATE: 1998-06-11
; PRIOR APPLICATION NUMBER: 60/089105
; PRIOR FILING DATE: 1998-06-12
; PRIOR APPLICATION NUMBER: 60/089440
; PRIOR FILING DATE: 1998-06-16
; PRIOR APPLICATION NUMBER: 60/089512
; PRIOR FILING DATE: 1998-06-16
; PRIOR APPLICATION NUMBER: 60/089514
; PRIOR FILING DATE: 1998-06-16
; PRIOR APPLICATION NUMBER: 60/089532
; PRIOR FILING DATE: 1998-06-17
; PRIOR APPLICATION NUMBER: 60/089538
; PRIOR FILING DATE: 1998-06-17
; PRIOR APPLICATION NUMBER: 60/089598
; PRIOR FILING DATE: 1998-06-17
; PRIOR APPLICATION NUMBER: 60/089599
; PRIOR FILING DATE: 1998-06-17
; PRIOR APPLICATION NUMBER: 60/089600
; PRIOR FILING DATE: 1998-06-17
; PRIOR APPLICATION NUMBER: 60/089653
; PRIOR FILING DATE: 1998-06-17
; PRIOR APPLICATION NUMBER: 60/089801
; PRIOR FILING DATE: 1998-06-18
; PRIOR APPLICATION NUMBER: 60/089907
; PRIOR FILING DATE: 1998-06-18
; PRIOR APPLICATION NUMBER: 60/089908
; PRIOR FILING DATE: 1998-06-18
; PRIOR APPLICATION NUMBER: 60/089947
; PRIOR FILING DATE: 1998-06-19
; PRIOR APPLICATION NUMBER: 60/089948
; PRIOR FILING DATE: 1998-06-19
; PRIOR APPLICATION NUMBER: 60/089952
; PRIOR FILING DATE: 1998-06-19
; PRIOR APPLICATION NUMBER: 60/090246
; PRIOR FILING DATE: 1998-06-22
; PRIOR APPLICATION NUMBER: 60/090252
; PRIOR FILING DATE: 1998-06-22
; PRIOR APPLICATION NUMBER: 60/090254
; PRIOR FILING DATE: 1998-06-22
; PRIOR APPLICATION NUMBER: 60/090349
; PRIOR FILING DATE: 1998-06-23
; PRIOR APPLICATION NUMBER: 60/090355
; PRIOR FILING DATE: 1998-06-23
; PRIOR APPLICATION NUMBER: 60/090429
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090431
; PRIOR FILING DATE: 1998-06-24

; PRIOR APPLICATION NUMBER: 60/090435
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090444
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090445
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090472
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090535
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090540
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090542
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090557
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090676
; PRIOR FILING DATE: 1998-06-25
; PRIOR APPLICATION NUMBER: 60/090678
; PRIOR FILING DATE: 1998-06-25
; PRIOR APPLICATION NUMBER: 60/090690
; PRIOR FILING DATE: 1998-06-25
; PRIOR APPLICATION NUMBER: 60/090694
; PRIOR FILING DATE: 1998-06-25
; PRIOR APPLICATION NUMBER: 60/090695
; PRIOR FILING DATE: 1998-06-25
; PRIOR APPLICATION NUMBER: 60/090696
; PRIOR FILING DATE: 1998-06-25
; PRIOR APPLICATION NUMBER: 60/090862
; PRIOR FILING DATE: 1998-06-26
; PRIOR APPLICATION NUMBER: 60/090863
; PRIOR FILING DATE: 1998-06-26
; PRIOR APPLICATION NUMBER: 60/091360
; PRIOR FILING DATE: 1998-07-01
; PRIOR APPLICATION NUMBER: 60/091478
; PRIOR FILING DATE: 1998-07-02
; PRIOR APPLICATION NUMBER: 60/091544
; PRIOR FILING DATE: 1998-07-01
; PRIOR APPLICATION NUMBER: 60/091519
; PRIOR FILING DATE: 1998-07-02
; PRIOR APPLICATION NUMBER: 60/091626
; PRIOR FILING DATE: 1998-07-02
; PRIOR APPLICATION NUMBER: 60/091633
; PRIOR FILING DATE: 1998-07-02
; PRIOR APPLICATION NUMBER: 60/091978
; PRIOR FILING DATE: 1998-07-07
; PRIOR APPLICATION NUMBER: 60/091982
; PRIOR FILING DATE: 1998-07-07
; PRIOR APPLICATION NUMBER: 60/092182
; PRIOR FILING DATE: 1998-07-09

Alignment Scores:

Pred. No.:	1.9e-05	Length:	762
Score:	112.00	Matches:	24
Percent Similarity:	53.57%	Conservative:	6
Best Local Similarity:	42.86%	Mismatches:	24
Query Match:	34.25%	Indels:	2
DB:	9	Gaps:	2

09-833799-13B (1-57) x US-09-990-436-344 (1-762)

QY	1	AlaGlnGluProValLysGlyProValSerThrLysProGlySerCysProIleLeu	20
Db	62	GCTGTGAAGGAGTTAAAGAGGGTATAGAG---AAAGCAGGGGTTTGGCCAGCTGACAAC	118
QY	21	IleArgCysAlaMetLeuAsnProAsnArgCysLeuLysAspThrAspCysProGly	40
Db	119	GTACGCTGCTTCAAGTCCGATCCTCCC---CAGTGTACACAGACCAGGACTGTCTGGG	175
QY	41	IleLysLysCysCysGluGlySerCysGlyMetAlaCysPheValPro	56
Db	176	GAAGGAAGTGTGTACCTGCTGCTGCTTCAAGTGTGTGATTCCT	223

Sun Feb 16 09:13:26 2003

09-833799-13b.rnpb

RESULT 14

US-09-991-181-344
; Sequence 344, Application US/09991181
; Publication No. US20020197615A1

; GENERAL INFORMATION:

; APPLICANT: Ashkenazi, Avi J.
; APPLICANT: Baker, Kevin P.
; APPLICANT: Botstein, David
; APPLICANT: Desnoyers, Luc
; APPLICANT: Eaton, Dan L.
; APPLICANT: Ferrara, Napoleone
; APPLICANT: Fong, Sherman
; APPLICANT: Gerber, Hanspeter
; APPLICANT: Gerritsen, Mary E.
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Grimaldi, J. Christopher
; APPLICANT: Gurney, Austin L.
; APPLICANT: Kljavin, Ivar J.
; APPLICANT: Napier, Mary A.
; APPLICANT: Pan, James
; APPLICANT: Paoni, Nicholas F.
; APPLICANT: Roy, Margaret Ann
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Watanabe, Colin K.
; APPLICANT: Williams, P. Mickey
; APPLICANT: Wood, William I.
; APPLICANT: Zhang, Zemin

; TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
; TITLE OF INVENTION: Acids Encoding the Same

; FILE REFERENCE: P2730P1C53

; CURRENT APPLICATION NUMBER: US/09/991,181

; CURRENT FILING DATE: 2001-11-16

; PRIOR APPLICATION NUMBER: 60/049787

; PRIOR FILING DATE: 1997-06-16

; PRIOR APPLICATION NUMBER: 60/062250

; PRIOR FILING DATE: 1997-10-17

; PRIOR APPLICATION NUMBER: 60/065186

; PRIOR FILING DATE: 1997-11-12

; PRIOR APPLICATION NUMBER: 60/065311

; PRIOR FILING DATE: 1997-11-13

; PRIOR APPLICATION NUMBER: 60/066770

; PRIOR FILING DATE: 1997-11-24

; PRIOR APPLICATION NUMBER: 60/075945

; PRIOR FILING DATE: 1998-02-25

; PRIOR APPLICATION NUMBER: 60/078910

; PRIOR FILING DATE: 1998-03-20

; PRIOR APPLICATION NUMBER: 60/083322

; PRIOR FILING DATE: 1998-04-28

; PRIOR APPLICATION NUMBER: 60/084600

; PRIOR FILING DATE: 1998-05-07

; PRIOR APPLICATION NUMBER: 60/087106

; PRIOR FILING DATE: 1998-05-28

; PRIOR APPLICATION NUMBER: 60/087607

; PRIOR FILING DATE: 1998-06-02

; PRIOR APPLICATION NUMBER: 60/087609

; PRIOR FILING DATE: 1998-06-02

; PRIOR APPLICATION NUMBER: 60/087759

; PRIOR FILING DATE: 1998-06-02

; PRIOR APPLICATION NUMBER: 60/087827

; PRIOR FILING DATE: 1998-06-03

; PRIOR APPLICATION NUMBER: 60/088021

; PRIOR FILING DATE: 1998-06-04

; PRIOR APPLICATION NUMBER: 60/088025

; PRIOR FILING DATE: 1998-06-04

; PRIOR APPLICATION NUMBER: 60/088026

; PRIOR FILING DATE: 1998-06-04

; PRIOR APPLICATION NUMBER: 60/088028

; PRIOR FILING DATE: 1998-06-04

; PRIOR APPLICATION NUMBER: 60/088029

; PRIOR FILING DATE: 1998-06-04

; PRIOR APPLICATION NUMBER: 60/088030

; PRIOR FILING DATE: 1998-06-23

; PRIOR FILING DATE: 1998-06-04
; PRIOR APPLICATION NUMBER: 60/088033
; PRIOR FILING DATE: 1998-06-04
; PRIOR APPLICATION NUMBER: 60/088326
; PRIOR FILING DATE: 1998-06-04
; PRIOR APPLICATION NUMBER: 60/088167
; PRIOR FILING DATE: 1998-06-05
; PRIOR APPLICATION NUMBER: 60/088202
; PRIOR FILING DATE: 1998-06-05
; PRIOR APPLICATION NUMBER: 60/088212
; PRIOR FILING DATE: 1998-06-05
; PRIOR APPLICATION NUMBER: 60/088217
; PRIOR FILING DATE: 1998-06-05
; PRIOR APPLICATION NUMBER: 60/088655
; PRIOR FILING DATE: 1998-06-09
; PRIOR APPLICATION NUMBER: 60/088734
; PRIOR FILING DATE: 1998-06-10
; PRIOR APPLICATION NUMBER: 60/088738
; PRIOR FILING DATE: 1998-06-10
; PRIOR APPLICATION NUMBER: 60/088742
; PRIOR FILING DATE: 1998-06-10
; PRIOR APPLICATION NUMBER: 60/088810
; PRIOR FILING DATE: 1998-06-10
; PRIOR APPLICATION NUMBER: 60/088824
; PRIOR FILING DATE: 1998-06-10
; PRIOR APPLICATION NUMBER: 60/088826
; PRIOR FILING DATE: 1998-06-10
; PRIOR APPLICATION NUMBER: 60/088858
; PRIOR FILING DATE: 1998-06-11
; PRIOR APPLICATION NUMBER: 60/088861
; PRIOR FILING DATE: 1998-06-11
; PRIOR APPLICATION NUMBER: 60/088876
; PRIOR FILING DATE: 1998-06-11
; PRIOR APPLICATION NUMBER: 60/089105
; PRIOR FILING DATE: 1998-06-12
; PRIOR APPLICATION NUMBER: 60/089440
; PRIOR FILING DATE: 1998-06-16
; PRIOR APPLICATION NUMBER: 60/089512
; PRIOR FILING DATE: 1998-06-16
; PRIOR APPLICATION NUMBER: 60/089514
; PRIOR FILING DATE: 1998-06-16
; PRIOR APPLICATION NUMBER: 60/089532
; PRIOR FILING DATE: 1998-06-17
; PRIOR APPLICATION NUMBER: 60/089538
; PRIOR FILING DATE: 1998-06-17
; PRIOR APPLICATION NUMBER: 60/089598
; PRIOR FILING DATE: 1998-06-17
; PRIOR APPLICATION NUMBER: 60/089599
; PRIOR FILING DATE: 1998-06-17
; PRIOR APPLICATION NUMBER: 60/089600
; PRIOR FILING DATE: 1998-06-17
; PRIOR APPLICATION NUMBER: 60/089653
; PRIOR FILING DATE: 1998-06-17
; PRIOR APPLICATION NUMBER: 60/089801
; PRIOR FILING DATE: 1998-06-18
; PRIOR APPLICATION NUMBER: 60/089907
; PRIOR FILING DATE: 1998-06-18
; PRIOR APPLICATION NUMBER: 60/089908
; PRIOR FILING DATE: 1998-06-18
; PRIOR APPLICATION NUMBER: 60/089947
; PRIOR FILING DATE: 1998-06-19
; PRIOR APPLICATION NUMBER: 60/089948
; PRIOR FILING DATE: 1998-06-19
; PRIOR APPLICATION NUMBER: 60/089952
; PRIOR FILING DATE: 1998-06-19
; PRIOR APPLICATION NUMBER: 60/090246
; PRIOR FILING DATE: 1998-06-22
; PRIOR APPLICATION NUMBER: 60/090252
; PRIOR FILING DATE: 1998-06-22
; PRIOR APPLICATION NUMBER: 60/090254
; PRIOR FILING DATE: 1998-06-22
; PRIOR APPLICATION NUMBER: 60/090349
; PRIOR FILING DATE: 1998-06-23

; PRIOR FILING DATE: 1998-06-04
; PRIOR APPLICATION NUMBER: 60/088028
; PRIOR FILING DATE: 1998-06-04
; PRIOR APPLICATION NUMBER: 60/088029
; PRIOR FILING DATE: 1998-06-04
; PRIOR APPLICATION NUMBER: 60/088030
; PRIOR FILING DATE: 1998-06-04
; PRIOR APPLICATION NUMBER: 60/088033
; PRIOR FILING DATE: 1998-06-04
; PRIOR APPLICATION NUMBER: 60/088326
; PRIOR FILING DATE: 1998-06-04
; PRIOR APPLICATION NUMBER: 60/088167
; PRIOR FILING DATE: 1998-06-05
; PRIOR APPLICATION NUMBER: 60/088202
; PRIOR FILING DATE: 1998-06-05
; PRIOR APPLICATION NUMBER: 60/088212
; PRIOR FILING DATE: 1998-06-05
; PRIOR APPLICATION NUMBER: 60/088217
; PRIOR FILING DATE: 1998-06-05
; PRIOR APPLICATION NUMBER: 60/088655
; PRIOR FILING DATE: 1998-06-09
; PRIOR APPLICATION NUMBER: 60/088734
; PRIOR FILING DATE: 1998-06-10
; PRIOR APPLICATION NUMBER: 60/088738
; PRIOR FILING DATE: 1998-06-10
; PRIOR APPLICATION NUMBER: 60/088742
; PRIOR FILING DATE: 1998-06-10
; PRIOR APPLICATION NUMBER: 60/088810
; PRIOR FILING DATE: 1998-06-10
; PRIOR APPLICATION NUMBER: 60/088824
; PRIOR FILING DATE: 1998-06-10
; PRIOR APPLICATION NUMBER: 60/088826
; PRIOR FILING DATE: 1998-06-10
; PRIOR APPLICATION NUMBER: 60/088858
; PRIOR FILING DATE: 1998-06-11
; PRIOR APPLICATION NUMBER: 60/088861
; PRIOR FILING DATE: 1998-06-11
; PRIOR APPLICATION NUMBER: 60/088876
; PRIOR FILING DATE: 1998-06-11
; PRIOR APPLICATION NUMBER: 60/089105
; PRIOR FILING DATE: 1998-06-12
; PRIOR APPLICATION NUMBER: 60/089440
; PRIOR FILING DATE: 1998-06-16
; PRIOR APPLICATION NUMBER: 60/089512
; PRIOR FILING DATE: 1998-06-16
; PRIOR APPLICATION NUMBER: 60/089514
; PRIOR FILING DATE: 1998-06-16
; PRIOR APPLICATION NUMBER: 60/089532
; PRIOR FILING DATE: 1998-06-17
; PRIOR APPLICATION NUMBER: 60/089538
; PRIOR FILING DATE: 1998-06-17
; PRIOR APPLICATION NUMBER: 60/089598
; PRIOR FILING DATE: 1998-06-17
; PRIOR APPLICATION NUMBER: 60/089599
; PRIOR FILING DATE: 1998-06-17
; PRIOR APPLICATION NUMBER: 60/089600
; PRIOR FILING DATE: 1998-06-17
; PRIOR APPLICATION NUMBER: 60/089653
; PRIOR FILING DATE: 1998-06-17
; PRIOR APPLICATION NUMBER: 60/089801
; PRIOR FILING DATE: 1998-06-18
; PRIOR APPLICATION NUMBER: 60/089907
; PRIOR FILING DATE: 1998-06-18
; PRIOR APPLICATION NUMBER: 60/089908
; PRIOR FILING DATE: 1998-06-18
; PRIOR APPLICATION NUMBER: 60/089947
; PRIOR FILING DATE: 1998-06-19
; PRIOR APPLICATION NUMBER: 60/089948
; PRIOR FILING DATE: 1998-06-19
; PRIOR APPLICATION NUMBER: 60/089952
; PRIOR FILING DATE: 1998-06-19
; PRIOR APPLICATION NUMBER: 60/090246
; PRIOR FILING DATE: 1998-06-22

; PRIOR APPLICATION NUMBER: 60/090252
; PRIOR FILING DATE: 1998-06-22
; PRIOR APPLICATION NUMBER: 60/090254
; PRIOR FILING DATE: 1998-06-22
; PRIOR APPLICATION NUMBER: 60/090349
; PRIOR FILING DATE: 1998-06-23
; PRIOR APPLICATION NUMBER: 60/090355
; PRIOR FILING DATE: 1998-06-23
; PRIOR APPLICATION NUMBER: 60/090429
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090431
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090435
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090444
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090445
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090472
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090535
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090540
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090542
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090557
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090676
; PRIOR FILING DATE: 1998-06-25
; PRIOR APPLICATION NUMBER: 60/090678
; PRIOR FILING DATE: 1998-06-25
; PRIOR APPLICATION NUMBER: 60/090690
; PRIOR FILING DATE: 1998-06-25
; PRIOR APPLICATION NUMBER: 60/090694
; PRIOR FILING DATE: 1998-06-25
; PRIOR APPLICATION NUMBER: 60/090695
; PRIOR FILING DATE: 1998-06-25
; PRIOR APPLICATION NUMBER: 60/090696
; PRIOR FILING DATE: 1998-06-25
; PRIOR APPLICATION NUMBER: 60/090862
; PRIOR FILING DATE: 1998-06-26
; PRIOR APPLICATION NUMBER: 60/090863
; PRIOR FILING DATE: 1998-06-26
; PRIOR APPLICATION NUMBER: 60/091360
; PRIOR FILING DATE: 1998-07-01
; PRIOR APPLICATION NUMBER: 60/091478
; PRIOR FILING DATE: 1998-07-02
; PRIOR APPLICATION NUMBER: 60/091544
; PRIOR FILING DATE: 1998-07-01
; PRIOR APPLICATION NUMBER: 60/091519
; PRIOR FILING DATE: 1998-07-02
; PRIOR APPLICATION NUMBER: 60/091626
; PRIOR FILING DATE: 1998-07-02
; PRIOR APPLICATION NUMBER: 60/091633
; PRIOR FILING DATE: 1998-07-02
; PRIOR APPLICATION NUMBER: 60/091978
; PRIOR FILING DATE: 1998-07-07
; PRIOR APPLICATION NUMBER: 60/091982
; PRIOR FILING DATE: 1998-07-07
; PRIOR APPLICATION NUMBER: 60/092182
; PRIOR FILING DATE: 1998-07-09

Alignment Scores:

Pred. No.:	1.9e-05	Length:	762
Score:	112.00	Matches:	24
Percent Similarity:	53.57%	Conservative:	6
Best Local Similarity:	42.86%	Mismatches:	24
Query Match:	34.25%	Indels:	2
DB:	9	Gaps:	2

09-833799-13B (1-57) x US-09-993-687-344 (1-762)

;	PRIOR	FILING DATE:	1998-06-03	
;	PRIOR	APPLICATION NUMBER:	60/088021	
;	PRIOR	FILING DATE:	1998-06-04	
;	PRIOR	APPLICATION NUMBER:	60/088025	
;	PRIOR	FILING DATE:	1998-06-04	
;	PRIOR	APPLICATION NUMBER:	60/088026	
;	PRIOR	FILING DATE:	1998-06-04	
;	PRIOR	APPLICATION NUMBER:	60/088028	
;	PRIOR	FILING DATE:	1998-06-04	
;	PRIOR	APPLICATION NUMBER:	60/088029	
;	PRIOR	FILING DATE:	1998-06-04	
;	PRIOR	APPLICATION NUMBER:	60/088030	
;	PRIOR	FILING DATE:	1998-06-04	
;	PRIOR	APPLICATION NUMBER:	60/088033	
;	PRIOR	FILING DATE:	1998-06-04	
;	PRIOR	APPLICATION NUMBER:	60/088326	
;	PRIOR	FILING DATE:	1998-06-04	
;	PRIOR	APPLICATION NUMBER:	60/088167	
;	PRIOR	FILING DATE:	1998-06-05	
;	PRIOR	APPLICATION NUMBER:	60/088202	
;	PRIOR	FILING DATE:	1998-06-05	
;	PRIOR	APPLICATION NUMBER:	60/088212	
;	PRIOR	FILING DATE:	1998-06-05	
;	PRIOR	APPLICATION NUMBER:	60/088217	
;	PRIOR	FILING DATE:	1998-06-05	
;	PRIOR	APPLICATION NUMBER:	60/088655	
;	PRIOR	FILING DATE:	1998-06-09	
;	PRIOR	APPLICATION NUMBER:	60/088734	
;	PRIOR	FILING DATE:	1998-06-10	
;	PRIOR	APPLICATION NUMBER:	60/088738	
;	PRIOR	FILING DATE:	1998-06-10	
;	PRIOR	APPLICATION NUMBER:	60/088742	
;	PRIOR	FILING DATE:	1998-06-10	
;	PRIOR	APPLICATION NUMBER:	60/088810	
;	PRIOR	FILING DATE:	1998-06-10	
;	PRIOR	APPLICATION NUMBER:	60/088824	
;	PRIOR	FILING DATE:	1998-06-10	
;	PRIOR	APPLICATION NUMBER:	60/088826	
;	PRIOR	FILING DATE:	1998-06-10	
;	PRIOR	APPLICATION NUMBER:	60/088858	
;	PRIOR	FILING DATE:	1998-06-11	
;	PRIOR	APPLICATION NUMBER:	60/088861	
;	PRIOR	FILING DATE:	1998-06-11	
;	PRIOR	APPLICATION NUMBER:	60/088876	
;	PRIOR	FILING DATE:	1998-06-11	
;	PRIOR	APPLICATION NUMBER:	60/089105	
;	PRIOR	FILING DATE:	1998-06-12	
;	PRIOR	APPLICATION NUMBER:	60/089440	
;	PRIOR	FILING DATE:	1998-06-16	
;	PRIOR	APPLICATION NUMBER:	60/089512	
;	PRIOR	FILING DATE:	1998-06-16	
;	PRIOR	APPLICATION NUMBER:	60/089514	
;	PRIOR	FILING DATE:	1998-06-16	
;	PRIOR	APPLICATION NUMBER:	60/089532	
;	PRIOR	FILING DATE:	1998-06-17	
;	PRIOR	APPLICATION NUMBER:	60/089538	
;	PRIOR	FILING DATE:	1998-06-17	
;	PRIOR	APPLICATION NUMBER:	60/089598	
;	PRIOR	FILING DATE:	1998-06-17	
;	PRIOR	APPLICATION NUMBER:	60/089599	
;	PRIOR	FILING DATE:	1998-06-17	
;	PRIOR	APPLICATION NUMBER:	60/089600	
;	PRIOR	FILING DATE:	1998-06-17	
;	PRIOR	APPLICATION NUMBER:	60/089653	
;	PRIOR	FILING DATE:	1998-06-17	
;	PRIOR	APPLICATION NUMBER:	60/089801	
;	PRIOR	FILING DATE:	1998-06-18	
;	PRIOR	APPLICATION NUMBER:	60/089907	
;	PRIOR	FILING DATE:	1998-06-18	
;	PRIOR	APPLICATION NUMBER:	60/089908	
;	PRIOR	FILING DATE:	1998-06-18	
;	PRIOR	APPLICATION NUMBER:	60/089947	
;	PRIOR	FILING DATE:	1998-06-19	

;	PRIOR APPLICATION NUMBER:	60/089900
;	PRIOR FILING DATE:	1998-06-18
;	PRIOR APPLICATION NUMBER:	60/089908
;	PRIOR FILING DATE:	1998-06-18
;	PRIOR APPLICATION NUMBER:	60/089947
;	PRIOR FILING DATE:	1998-06-19
;	PRIOR APPLICATION NUMBER:	60/089948
;	PRIOR FILING DATE:	1998-06-19
;	PRIOR APPLICATION NUMBER:	60/089952
;	PRIOR FILING DATE:	1998-06-19
;	PRIOR APPLICATION NUMBER:	60/090246
;	PRIOR FILING DATE:	1998-06-22
;	PRIOR APPLICATION NUMBER:	60/090252
;	PRIOR FILING DATE:	1998-06-22
;	PRIOR APPLICATION NUMBER:	60/090254
;	PRIOR FILING DATE:	1998-06-22
;	PRIOR APPLICATION NUMBER:	60/090349
;	PRIOR FILING DATE:	1998-06-23
;	PRIOR APPLICATION NUMBER:	60/090355
;	PRIOR FILING DATE:	1998-06-23
;	PRIOR APPLICATION NUMBER:	60/090429
;	PRIOR FILING DATE:	1998-06-24
;	PRIOR APPLICATION NUMBER:	60/090431
;	PRIOR FILING DATE:	1998-06-24
;	PRIOR APPLICATION NUMBER:	60/090435
;	PRIOR FILING DATE:	1998-06-24
;	PRIOR APPLICATION NUMBER:	60/090444
;	PRIOR FILING DATE:	1998-06-24
;	PRIOR APPLICATION NUMBER:	60/090445
;	PRIOR FILING DATE:	1998-06-24
;	PRIOR APPLICATION NUMBER:	60/090472
;	PRIOR FILING DATE:	1998-06-24
;	PRIOR APPLICATION NUMBER:	60/090535
;	PRIOR FILING DATE:	1998-06-24
;	PRIOR APPLICATION NUMBER:	60/090540
;	PRIOR FILING DATE:	1998-06-24
;	PRIOR APPLICATION NUMBER:	60/090542
;	PRIOR FILING DATE:	1998-06-24
;	PRIOR APPLICATION NUMBER:	60/090557
;	PRIOR FILING DATE:	1998-06-24
;	PRIOR APPLICATION NUMBER:	60/090676
;	PRIOR FILING DATE:	1998-06-25
;	PRIOR APPLICATION NUMBER:	60/090678
;	PRIOR FILING DATE:	1998-06-25
;	PRIOR APPLICATION NUMBER:	60/090690
;	PRIOR FILING DATE:	1998-06-25
;	PRIOR APPLICATION NUMBER:	60/090694
;	PRIOR FILING DATE:	1998-06-25
;	PRIOR APPLICATION NUMBER:	60/090695
;	PRIOR FILING DATE:	1998-06-25
;	PRIOR APPLICATION NUMBER:	60/090696
;	PRIOR FILING DATE:	1998-06-25
;	PRIOR APPLICATION NUMBER:	60/090862
;	PRIOR FILING DATE:	1998-06-26
;	PRIOR APPLICATION NUMBER:	60/090863
;	PRIOR FILING DATE:	1998-06-26
;	PRIOR APPLICATION NUMBER:	60/091360
;	PRIOR FILING DATE:	1998-07-01
;	PRIOR APPLICATION NUMBER:	60/091478
;	PRIOR FILING DATE:	1998-07-02
;	PRIOR APPLICATION NUMBER:	60/091544
;	PRIOR FILING DATE:	1998-07-01
;	PRIOR APPLICATION NUMBER:	60/091519
;	PRIOR FILING DATE:	1998-07-02
;	PRIOR APPLICATION NUMBER:	60/091626
;	PRIOR FILING DATE:	1998-07-02
;	PRIOR APPLICATION NUMBER:	60/091633
;	PRIOR FILING DATE:	1998-07-02
;	PRIOR APPLICATION NUMBER:	60/091978
;	PRIOR FILING DATE:	1998-07-07
;	PRIOR APPLICATION NUMBER:	60/091982
;	PRIOR FILING DATE:	1998-07-07
;	PRIOR APPLICATION NUMBER:	60/092182

09-833799-13b.rnpb

Sun Feb 16 09:13:26 2003

; PRIOR FILING DATE: 1998-07-09

Alignment Scores: 1.9e-05 762
Pred. No.: 112.00 24
Score: 53.57% 6
Percent Similarity: 42.86% 24
Best Local Similarity: 34.25% 2
Query Match: 9 2
DB:

09-833799-13B (1-57) x US-09-997-653-344 (1-762)

Qy 1 AlaGlnGluProValLysGlyProValSerThrLysProGlySerCysProIleLeu 20
Db 62 GCTGTGAAGGAGTTAAAGAGGGTATAGAG---AAAGCAGGGGTTTCCCCAGCTGACAAC 118
Qy 21 IleArgCysAlaMetLeuAsnProProAsnArgCysLeuLysAspThrAspCysProGly 40
Db 119 GTACGCTGCTTCAAGTCCGATCCTCCC---CAGTGTACACACAGACCAGGACTGTCTGGGG 175
Qy 41 IleLysLysCysCysGluGlySerCysGlyMetAlaCysPheValPro 56
Db 176 GAAAGGAAGTGTGTTACCTGCACCTGTGGCTTCAAGTGTGTGATTCT 223

RESULT 18

US-09-993-667-344 Application US/09993667
; Publication No. US20030022187A1

GENERAL INFORMATION:

; APPLICANT: Ashkenazi, Avi J.
; APPLICANT: Baker, Kevin P.
; APPLICANT: Botstein, David
; APPLICANT: Desnoyers, Luc
; APPLICANT: Eaton, Dan L.
; APPLICANT: Ferrara, Napoleone
; APPLICANT: Fong, Sherman
; APPLICANT: Gerber, Hanspeter
; APPLICANT: Gerritsen, Mary E.
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Grimaldi, J. Christopher
; APPLICANT: Gurney, Austin L.
; APPLICANT: Kljavin, Ivar J.
; APPLICANT: Napier, Mary A.
; APPLICANT: Pan, James
; APPLICANT: Paoni, Nicholas F.
; APPLICANT: Roy, Margaret Ann
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Watanabe, Colin K.
; APPLICANT: Williams, P. Mickey
; APPLICANT: Wood, William I.
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
; FILE REFERENCE: P2730P1C4
; CURRENT APPLICATION NUMBER: US/09/993,667
; CURRENT FILING DATE: 2001-11-14
; PRIOR APPLICATION NUMBER: 60/049787
; PRIOR FILING DATE: 1997-06-16
; PRIOR APPLICATION NUMBER: 60/062250
; PRIOR FILING DATE: 1997-10-17
; PRIOR APPLICATION NUMBER: 60/065186
; PRIOR FILING DATE: 1997-11-12
; PRIOR APPLICATION NUMBER: 60/065311
; PRIOR FILING DATE: 1997-11-13
; PRIOR APPLICATION NUMBER: 60/066770
; PRIOR FILING DATE: 1997-11-24
; PRIOR APPLICATION NUMBER: 60/075945
; PRIOR FILING DATE: 1998-02-25
; PRIOR APPLICATION NUMBER: 60/078910
; PRIOR FILING DATE: 1998-03-20
; PRIOR APPLICATION NUMBER: 60/083322

; PRIOR FILING DATE: 1998-04-28
; PRIOR APPLICATION NUMBER: 60/084600
; PRIOR FILING DATE: 1998-05-07
; PRIOR APPLICATION NUMBER: 60/087106
; PRIOR FILING DATE: 1998-05-28
; PRIOR APPLICATION NUMBER: 60/087607
; PRIOR FILING DATE: 1998-06-02
; PRIOR APPLICATION NUMBER: 60/087609
; PRIOR FILING DATE: 1998-06-02
; PRIOR APPLICATION NUMBER: 60/087759
; PRIOR FILING DATE: 1998-06-02
; PRIOR APPLICATION NUMBER: 60/087827
; PRIOR FILING DATE: 1998-06-03
; PRIOR APPLICATION NUMBER: 60/088021
; PRIOR FILING DATE: 1998-06-04
; PRIOR APPLICATION NUMBER: 60/088025
; PRIOR FILING DATE: 1998-06-04
; PRIOR APPLICATION NUMBER: 60/088026
; PRIOR FILING DATE: 1998-06-04
; PRIOR APPLICATION NUMBER: 60/088028
; PRIOR FILING DATE: 1998-06-04
; PRIOR APPLICATION NUMBER: 60/088029
; PRIOR FILING DATE: 1998-06-04
; PRIOR APPLICATION NUMBER: 60/088030
; PRIOR FILING DATE: 1998-06-04
; PRIOR APPLICATION NUMBER: 60/088033
; PRIOR FILING DATE: 1998-06-04
; PRIOR APPLICATION NUMBER: 60/088326
; PRIOR FILING DATE: 1998-06-04
; PRIOR APPLICATION NUMBER: 60/088167
; PRIOR FILING DATE: 1998-06-05
; PRIOR APPLICATION NUMBER: 60/088202
; PRIOR FILING DATE: 1998-06-05
; PRIOR APPLICATION NUMBER: 60/088212
; PRIOR FILING DATE: 1998-06-05
; PRIOR APPLICATION NUMBER: 60/088217
; PRIOR FILING DATE: 1998-06-05
; PRIOR APPLICATION NUMBER: 60/088655
; PRIOR FILING DATE: 1998-06-09
; PRIOR APPLICATION NUMBER: 60/088734
; PRIOR FILING DATE: 1998-06-10
; PRIOR APPLICATION NUMBER: 60/088738
; PRIOR FILING DATE: 1998-06-10
; PRIOR APPLICATION NUMBER: 60/088742
; PRIOR FILING DATE: 1998-06-10
; PRIOR APPLICATION NUMBER: 60/088810
; PRIOR FILING DATE: 1998-06-10
; PRIOR APPLICATION NUMBER: 60/088824
; PRIOR FILING DATE: 1998-06-10
; PRIOR APPLICATION NUMBER: 60/088826
; PRIOR FILING DATE: 1998-06-10
; PRIOR APPLICATION NUMBER: 60/088858
; PRIOR FILING DATE: 1998-06-11
; PRIOR APPLICATION NUMBER: 60/088861
; PRIOR FILING DATE: 1998-06-11
; PRIOR APPLICATION NUMBER: 60/088876
; PRIOR FILING DATE: 1998-06-11
; PRIOR APPLICATION NUMBER: 60/089105
; PRIOR FILING DATE: 1998-06-12
; PRIOR APPLICATION NUMBER: 60/089440
; PRIOR FILING DATE: 1998-06-16
; PRIOR APPLICATION NUMBER: 60/089512
; PRIOR FILING DATE: 1998-06-16
; PRIOR APPLICATION NUMBER: 60/089514
; PRIOR FILING DATE: 1998-06-16
; PRIOR APPLICATION NUMBER: 60/089532
; PRIOR FILING DATE: 1998-06-17
; PRIOR APPLICATION NUMBER: 60/089538
; PRIOR FILING DATE: 1998-06-17
; PRIOR APPLICATION NUMBER: 60/089598
; PRIOR FILING DATE: 1998-06-17
; PRIOR APPLICATION NUMBER: 60/089599
; PRIOR FILING DATE: 1998-06-17

; PRIOR APPLICATION NUMBER: 60/089600
; PRIOR FILING DATE: 1998-06-17
; PRIOR APPLICATION NUMBER: 60/089653
; PRIOR FILING DATE: 1998-06-17
; PRIOR APPLICATION NUMBER: 60/089801
; PRIOR FILING DATE: 1998-06-18
; PRIOR APPLICATION NUMBER: 60/089907
; PRIOR FILING DATE: 1998-06-18
; PRIOR APPLICATION NUMBER: 60/089908
; PRIOR FILING DATE: 1998-06-18
; PRIOR APPLICATION NUMBER: 60/089947
; PRIOR FILING DATE: 1998-06-19
; PRIOR APPLICATION NUMBER: 60/089948
; PRIOR FILING DATE: 1998-06-19
; PRIOR APPLICATION NUMBER: 60/089952
; PRIOR FILING DATE: 1998-06-19
; PRIOR APPLICATION NUMBER: 60/090246
; PRIOR FILING DATE: 1998-06-22
; PRIOR APPLICATION NUMBER: 60/090252
; PRIOR FILING DATE: 1998-06-22
; PRIOR APPLICATION NUMBER: 60/090254
; PRIOR FILING DATE: 1998-06-22
; PRIOR APPLICATION NUMBER: 60/090349
; PRIOR FILING DATE: 1998-06-23
; PRIOR APPLICATION NUMBER: 60/090355
; PRIOR FILING DATE: 1998-06-23
; PRIOR APPLICATION NUMBER: 60/090429
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090431
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090435
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090444
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090445
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090472
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090535
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090540
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090542
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090557
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090676
; PRIOR FILING DATE: 1998-06-25
; PRIOR APPLICATION NUMBER: 60/090678
; PRIOR FILING DATE: 1998-06-25
; PRIOR APPLICATION NUMBER: 60/090690
; PRIOR FILING DATE: 1998-06-25
; PRIOR APPLICATION NUMBER: 60/090694
; PRIOR FILING DATE: 1998-06-25
; PRIOR APPLICATION NUMBER: 60/090695
; PRIOR FILING DATE: 1998-06-25
; PRIOR APPLICATION NUMBER: 60/090696
; PRIOR FILING DATE: 1998-06-25
; PRIOR APPLICATION NUMBER: 60/090862
; PRIOR FILING DATE: 1998-06-26
; PRIOR APPLICATION NUMBER: 60/090863
; PRIOR FILING DATE: 1998-06-26
; PRIOR APPLICATION NUMBER: 60/091360
; PRIOR FILING DATE: 1998-07-01
; PRIOR APPLICATION NUMBER: 60/091478
; PRIOR FILING DATE: 1998-07-02
; PRIOR APPLICATION NUMBER: 60/091544
; PRIOR FILING DATE: 1998-07-01
; PRIOR APPLICATION NUMBER: 60/091519
; PRIOR FILING DATE: 1998-07-02
; PRIOR APPLICATION NUMBER: 60/091626
; PRIOR FILING DATE: 1998-07-02
; PRIOR APPLICATION NUMBER: 60/091633

; PRIOR FILING DATE: 1998-07-02
; PRIOR APPLICATION NUMBER: 60/091978
; PRIOR FILING DATE: 1998-07-07
; PRIOR APPLICATION NUMBER: 60/091982
; PRIOR FILING DATE: 1998-07-07
; PRIOR APPLICATION NUMBER: 60/092182
; PRIOR FILING DATE: 1998-07-09

Alignment Scores:
Pred. No.: 1.9e-05 Length: 762
Score: 112.00 Matches: 24
Percent Similarity: 53.57% Conservative: 6
Best Local Similarity: 42.86% Mismatches: 24
Query Match: 34.25% Indels: 2
DB: 9 Gaps: 2

09-833799-13B (1-57) x US-09-993-667-344 (1-762)

QY 1 AlaGlnGluProValLysGlyProValSerThrLysProGlySerCysProIleIleLeu 20
Db 62 GCTGTGAAGGAGTTAAAGAGGGTATAGAG---AAAGCAGGGGTTTGGCCAGCTGACAAAC 118

QY 21 IleArgCysAlaMetLeuAsnProProAsnArgCysLeuLysAspThrAspCysProGly 40
Db 119 GTACGCTGCTTCAAGTCCGATCCTCCC--CAGTGTACACACAGACCAGGACTGTCTGGGG 175

QY 41 IleLysLysCysCysGluGlySerCysGlyMetAlaCysPheValPro 56
Db 176 GAAAGGAAGTGTGTGTACCTGCACTGTGGCTTCAAGTGTGTGATTCCT 223

RESULT 19
US-09-990-438-344
; Sequence 344, Application US/09990438
; Publication No. US20030027754A1
; GENERAL INFORMATION:
; APPLICANT: Ashkenazi, Avi J.
; APPLICANT: Baker, Kevin P.
; APPLICANT: Botstein, David
; APPLICANT: Desnoyers, Luc
; APPLICANT: Eaton, Dan L.
; APPLICANT: Ferrara, Napoleone
; APPLICANT: Fong, Sherman
; APPLICANT: Gerber, Hanspeter
; APPLICANT: Gerritsen, Mary E.
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Grimaldi, J. Christopher
; APPLICANT: Gurney, Austin L.
; APPLICANT: Kljavin, Ivar J.
; APPLICANT: Napier, Mary A.
; APPLICANT: Pan, James
; APPLICANT: Paoni, Nicholas F.
; APPLICANT: Roy, Margaret Ann
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Watanabe, Colin K.
; APPLICANT: Williams, P. Mickey
; APPLICANT: Wood, William I.
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
; FILE OF INVENTION: Acids Encoding the Same
; FILE REFERENCE: P2730P1C3
; CURRENT APPLICATION NUMBER: US/09/990,438
; CURRENT FILING DATE: 2001-11-14
; PRIOR APPLICATION NUMBER: 60/049787
; PRIOR FILING DATE: 1997-06-16
; PRIOR APPLICATION NUMBER: 60/062250
; PRIOR FILING DATE: 1997-10-17
; PRIOR APPLICATION NUMBER: 60/065186
; PRIOR FILING DATE: 1997-11-12
; PRIOR APPLICATION NUMBER: 60/065311
; PRIOR FILING DATE: 1997-11-13
; PRIOR APPLICATION NUMBER: 60/066770

;	PRIOR FILING DATE:	1997-10-17	
;	PRIOR APPLICATION NUMBER:	60/065188	
;	PRIOR FILING DATE:	1997-11-12	
;	PRIOR APPLICATION NUMBER:	60/065311	
;	PRIOR FILING DATE:	1997-11-13	
;	PRIOR APPLICATION NUMBER:	60/066770	
;	PRIOR FILING DATE:	1997-11-24	
;	PRIOR APPLICATION NUMBER:	60/075945	
;	PRIOR FILING DATE:	1998-02-25	
;	PRIOR APPLICATION NUMBER:	60/078910	
;	PRIOR FILING DATE:	1998-03-20	
;	PRIOR APPLICATION NUMBER:	60/083322	
;	PRIOR FILING DATE:	1998-04-28	
;	PRIOR APPLICATION NUMBER:	60/084600	
;	PRIOR FILING DATE:	1998-05-07	
;	PRIOR APPLICATION NUMBER:	60/087106	
;	PRIOR FILING DATE:	1998-05-28	
;	PRIOR APPLICATION NUMBER:	60/087607	
;	PRIOR FILING DATE:	1998-06-02	
;	PRIOR APPLICATION NUMBER:	60/087609	
;	PRIOR FILING DATE:	1998-06-02	
;	PRIOR APPLICATION NUMBER:	60/087759	
;	PRIOR FILING DATE:	1998-06-02	
;	PRIOR APPLICATION NUMBER:	60/087827	
;	PRIOR FILING DATE:	1998-06-03	
;	PRIOR APPLICATION NUMBER:	60/088021	
;	PRIOR FILING DATE:	1998-06-04	
;	PRIOR APPLICATION NUMBER:	60/088025	
;	PRIOR FILING DATE:	1998-06-04	
;	PRIOR APPLICATION NUMBER:	60/088026	
;	PRIOR FILING DATE:	1998-06-04	
;	PRIOR APPLICATION NUMBER:	60/088028	
;	PRIOR FILING DATE:	1998-06-04	
;	PRIOR APPLICATION NUMBER:	60/088029	
;	PRIOR FILING DATE:	1998-06-04	
;	PRIOR APPLICATION NUMBER:	60/088030	
;	PRIOR FILING DATE:	1998-06-04	
;	PRIOR APPLICATION NUMBER:	60/088033	
;	PRIOR FILING DATE:	1998-06-04	
;	PRIOR APPLICATION NUMBER:	60/088326	
;	PRIOR FILING DATE:	1998-06-04	
;	PRIOR APPLICATION NUMBER:	60/088167	
;	PRIOR FILING DATE:	1998-06-05	
;	PRIOR APPLICATION NUMBER:	60/088202	
;	PRIOR FILING DATE:	1998-06-05	
;	PRIOR APPLICATION NUMBER:	60/088212	
;	PRIOR FILING DATE:	1998-06-05	
;	PRIOR APPLICATION NUMBER:	60/088217	
;	PRIOR FILING DATE:	1998-06-05	
;	PRIOR APPLICATION NUMBER:	60/088655	
;	PRIOR FILING DATE:	1998-06-09	
;	PRIOR APPLICATION NUMBER:	60/088734	
;	PRIOR FILING DATE:	1998-06-10	
;	PRIOR APPLICATION NUMBER:	60/088738	
;	PRIOR FILING DATE:	1998-06-10	
;	PRIOR APPLICATION NUMBER:	60/088742	
;	PRIOR FILING DATE:	1998-06-10	
;	PRIOR APPLICATION NUMBER:	60/088810	
;	PRIOR FILING DATE:	1998-06-10	
;	PRIOR APPLICATION NUMBER:	60/088824	
;	PRIOR FILING DATE:	1998-06-10	
;	PRIOR APPLICATION NUMBER:	60/088826	
;	PRIOR FILING DATE:	1998-06-10	
;	PRIOR APPLICATION NUMBER:	60/088858	
;	PRIOR FILING DATE:	1998-06-11	
;	PRIOR APPLICATION NUMBER:	60/088861	
;	PRIOR FILING DATE:	1998-06-11	
;	PRIOR APPLICATION NUMBER:	60/088876	
;	PRIOR FILING DATE:	1998-06-11	
;	PRIOR APPLICATION NUMBER:	60/089105	
;	PRIOR FILING DATE:	1998-06-12	
;	PRIOR APPLICATION NUMBER:	60/089440	
;	PRIOR FILING DATE:	1998-06-16	

Alignment Scores:	
Pred. No.:	1.9e-05
Score:	112.00
Percent Similarity:	53.57%
Best Local Similarity:	42.86%
Query Match:	34.25%
DB:	.
	9
Length:	762
Matches:	24
Conservative:	6
Mismatches:	24
Indels:	2
Gaps:	2

Qy	1	AlaGlnGluProValLysGlyProValSerThrLysProGlySerCysProIleIleLeu	20
		::	
Db	62	GCTGTGGAAGGAGTTAAAGAGGGGTATAGAG---AAAGCAGGGGTTTGGCCAGCTGACAAC	118
Qy	21	IleArgCysAlaMetLeuAsnProProAsnArgCysLeuLysAspThrAspCysProGly	40
		::: ::: :::	
Db	119	GTACGGTGTCTTCAAGTCCGATCTCTCCC---CAGTGTCAACACAGACCAGGACTGTCTGGGG	175

Db 176 GAAAGGAAGTGTGTTACCTGCACCTGTGGCTTCAAGTGTGATTCCCT 223

US-09-990-562-344

APPLICANT: Ashkenazi, Avi J.
APPLICANT: Baker, Kevin P.
APPLICANT: Botstein, David
APPLICANT: Desnoyers, Luc
APPLICANT: Eaton, Dan L.
APPLICANT: Ferrara, Napoleone
APPLICANT: Fong, Sherman
APPLICANT: Gerber, Hanspeter
APPLICANT: Gerritsen, Mary E.
APPLICANT: Goddard, Audrey
APPLICANT: Godowski, Paul J.
APPLICANT: Grimaldi, J. Christopher
APPLICANT: Gurney, Austin L.
APPLICANT: Kljavin, Ivar J.
APPLICANT: Napier, Mary A.
APPLICANT: Pan, James
APPLICANT: Paoni, Nicholas F.
APPLICANT: Rooy, Margaret Ann
APPLICANT: Stewart, Timothy A.
APPLICANT: Tumas, Daniel
APPLICANT: Watanabe, Colin K.
APPLICANT: Williams, P. Mickey
APPLICANT: Wood, William I.
APPLICANT: Zhang, Zemin
TITLE OF INVENTION: Secrets and Tr
TITLE OF INVENTION: Acids Encoding
FILE REFERENCE: P2730PIC18
CURRENT APPLICATION NUMBER: US/09/9
CURRENT FILING DATE: 2001-11-14
PRIOR APPLICATION NUMBER: 60/049787
PRIOR FILING DATE: 1997-06-16
PRIOR APPLICATION NUMBER: 60 062250

; PRIOR APPLICATION NUMBER: 60/089512
; PRIOR FILING DATE: 1998-06-16
; PRIOR APPLICATION NUMBER: 60/089514
; PRIOR FILING DATE: 1998-06-16
; PRIOR APPLICATION NUMBER: 60/089532
; PRIOR FILING DATE: 1998-06-17
; PRIOR APPLICATION NUMBER: 60/089538
; PRIOR FILING DATE: 1998-06-17
; PRIOR APPLICATION NUMBER: 60/089598
; PRIOR FILING DATE: 1998-06-17
; PRIOR APPLICATION NUMBER: 60/089599
; PRIOR FILING DATE: 1998-06-17
; PRIOR APPLICATION NUMBER: 60/089600
; PRIOR FILING DATE: 1998-06-17
; PRIOR APPLICATION NUMBER: 60/089653
; PRIOR FILING DATE: 1998-06-17
; PRIOR APPLICATION NUMBER: 60/089801
; PRIOR FILING DATE: 1998-06-18
; PRIOR APPLICATION NUMBER: 60/089907
; PRIOR FILING DATE: 1998-06-18
; PRIOR APPLICATION NUMBER: 60/089908
; PRIOR FILING DATE: 1998-06-18
; PRIOR APPLICATION NUMBER: 60/089947
; PRIOR FILING DATE: 1998-06-19
; PRIOR APPLICATION NUMBER: 60/089948
; PRIOR FILING DATE: 1998-06-19
; PRIOR APPLICATION NUMBER: 60/089952
; PRIOR FILING DATE: 1998-06-19
; PRIOR APPLICATION NUMBER: 60/090246
; PRIOR FILING DATE: 1998-06-22
; PRIOR APPLICATION NUMBER: 60/090252
; PRIOR FILING DATE: 1998-06-22
; PRIOR APPLICATION NUMBER: 60/090254
; PRIOR FILING DATE: 1998-06-22
; PRIOR APPLICATION NUMBER: 60/090349
; PRIOR FILING DATE: 1998-06-23
; PRIOR APPLICATION NUMBER: 60/090355
; PRIOR FILING DATE: 1998-06-23
; PRIOR APPLICATION NUMBER: 60/090429
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090431
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090435
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090444
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090445
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090472
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090535
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090540
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090542
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090557
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090676
; PRIOR FILING DATE: 1998-06-25
; PRIOR APPLICATION NUMBER: 60/090678
; PRIOR FILING DATE: 1998-06-25
; PRIOR APPLICATION NUMBER: 60/090690
; PRIOR FILING DATE: 1998-06-25
; PRIOR APPLICATION NUMBER: 60/090694
; PRIOR FILING DATE: 1998-06-25
; PRIOR APPLICATION NUMBER: 60/090695
; PRIOR FILING DATE: 1998-06-25
; PRIOR APPLICATION NUMBER: 60/090696
; PRIOR FILING DATE: 1998-06-25
; PRIOR APPLICATION NUMBER: 60/090862
; PRIOR FILING DATE: 1998-06-26
; PRIOR APPLICATION NUMBER: 60/090863

; PRIOR FILING DATE: 1998-06-26
; PRIOR APPLICATION NUMBER: 60/091360
; PRIOR FILING DATE: 1998-07-01
; PRIOR APPLICATION NUMBER: 60/091478
; PRIOR FILING DATE: 1998-07-02
; PRIOR APPLICATION NUMBER: 60/091544
; PRIOR FILING DATE: 1998-07-01
; PRIOR APPLICATION NUMBER: 60/091519
; PRIOR FILING DATE: 1998-07-02
; PRIOR APPLICATION NUMBER: 60/091626
; PRIOR FILING DATE: 1998-07-02
; PRIOR APPLICATION NUMBER: 60/091633
; PRIOR FILING DATE: 1998-07-02
; PRIOR APPLICATION NUMBER: 60/091978
; PRIOR FILING DATE: 1998-07-07
; PRIOR APPLICATION NUMBER: 60/091982
; PRIOR FILING DATE: 1998-07-07
; PRIOR APPLICATION NUMBER: 60/092182
; PRIOR FILING DATE: 1998-07-09

Alignment Scores:
Pred. No.: 1.9e-05 Length: 762
Score: 112.00 Matches: 24
Percent Similarity: 53.57% Conservative: 6
Best Local Similarity: 42.86% Mismatches: 24
Query Match: 34.25% Indels: 2
DB: 9 Gaps: 2

09-833799-13B (1-57) x US-09-990-562-344 (1-762)

Qy 1 AlaGlnGluProValLysGlyProValSerThrLysProGlySerCysProIleLeu 20
Db 62 GCTGTGAAGGAGTTAAAGAGGGTATAGAG---AAAGCAGGGGTTTGGCCAGCTGACAAC 118
Qy 21 IleArgCysAlaMetLeuAsnProProAsnArgCysLeuLysAspThrAspCysProGly 40
Db 119 GTACGCTGCTTCAAGTCCGATCTCTCCC---CAGTGTCAACAGACAGGAGCTGTCTGGG 175
Qy 41 IleLysLysCysCysGluGlySerCysGlyMetAlaCysPheValPro 56
Db 176 GAAAGGAAGTGTGTTACCTGCACCTGTGGCTTCAAGTGTGTGATTCTT 223

RESULT 21
US-09-997-428-344
; Sequence 344, Application US/09997428
; Publication No. US20030027162A1
; GENERAL INFORMATION:
; APPLICANT: Ashkenazi, Avi J.
; APPLICANT: Baker, Kevin P.
; APPLICANT: Botstein, David
; APPLICANT: Desnoyers, Luc
; APPLICANT: Eaton, Dan L.
; APPLICANT: Ferrara, Napoleone
; APPLICANT: Fong, Sherman
; APPLICANT: Gerber, Hanspeter
; APPLICANT: Gerritsen, Mary E.
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Grimaldi, J. Christopher
; APPLICANT: Gurney, Austin L.
; APPLICANT: Kljavin, Ivar J.
; APPLICANT: Napier, Mary A.
; APPLICANT: Pan, James
; APPLICANT: Paoni, Nicholas F.
; APPLICANT: Roy, Margaret Ann
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Watanabe, Colin K.
; APPLICANT: Williams, P. Mickey
; APPLICANT: Wood, William I.
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
; TITLE OF INVENTION: Acids Encoding the Same

;	PRIOR APPLICATION NUMBER: 60/088870
;	PRIOR FILING DATE: 1998-06-11
;	PRIOR APPLICATION NUMBER: 60/089105
;	PRIOR FILING DATE: 1998-06-12
;	PRIOR APPLICATION NUMBER: 60/089440
;	PRIOR FILING DATE: 1998-06-16
;	PRIOR APPLICATION NUMBER: 60/089512
;	PRIOR FILING DATE: 1998-06-16
;	PRIOR APPLICATION NUMBER: 60/089514
;	PRIOR FILING DATE: 1998-06-16
;	PRIOR APPLICATION NUMBER: 60/089532
;	PRIOR FILING DATE: 1998-06-17
;	PRIOR APPLICATION NUMBER: 60/089538
;	PRIOR FILING DATE: 1998-06-17
;	PRIOR APPLICATION NUMBER: 60/089598
;	PRIOR FILING DATE: 1998-06-17
;	PRIOR APPLICATION NUMBER: 60/089599
;	PRIOR FILING DATE: 1998-06-17
;	PRIOR APPLICATION NUMBER: 60/089600
;	PRIOR FILING DATE: 1998-06-17
;	PRIOR APPLICATION NUMBER: 60/089653
;	PRIOR FILING DATE: 1998-06-17
;	PRIOR APPLICATION NUMBER: 60/089801
;	PRIOR FILING DATE: 1998-06-18
;	PRIOR APPLICATION NUMBER: 60/089907
;	PRIOR FILING DATE: 1998-06-18
;	PRIOR APPLICATION NUMBER: 60/089908
;	PRIOR FILING DATE: 1998-06-18
;	PRIOR APPLICATION NUMBER: 60/089947
;	PRIOR FILING DATE: 1998-06-19
;	PRIOR APPLICATION NUMBER: 60/089948
;	PRIOR FILING DATE: 1998-06-19
;	PRIOR APPLICATION NUMBER: 60/089952
;	PRIOR FILING DATE: 1998-06-19
;	PRIOR APPLICATION NUMBER: 60/090246
;	PRIOR FILING DATE: 1998-06-22
;	PRIOR APPLICATION NUMBER: 60/090252
;	PRIOR FILING DATE: 1998-06-22
;	PRIOR APPLICATION NUMBER: 60/090254
;	PRIOR FILING DATE: 1998-06-22
;	PRIOR APPLICATION NUMBER: 60/090349
;	PRIOR FILING DATE: 1998-06-23
;	PRIOR APPLICATION NUMBER: 60/090355
;	PRIOR FILING DATE: 1998-06-23
;	PRIOR APPLICATION NUMBER: 60/090429
;	PRIOR FILING DATE: 1998-06-24
;	PRIOR APPLICATION NUMBER: 60/090431
;	PRIOR FILING DATE: 1998-06-24
;	PRIOR APPLICATION NUMBER: 60/090435
;	PRIOR FILING DATE: 1998-06-24
;	PRIOR APPLICATION NUMBER: 60/090444
;	PRIOR FILING DATE: 1998-06-24
;	PRIOR APPLICATION NUMBER: 60/090445
;	PRIOR FILING DATE: 1998-06-24
;	PRIOR APPLICATION NUMBER: 60/090472
;	PRIOR FILING DATE: 1998-06-24
;	PRIOR APPLICATION NUMBER: 60/090535
;	PRIOR FILING DATE: 1998-06-24
;	PRIOR APPLICATION NUMBER: 60/090540
;	PRIOR FILING DATE: 1998-06-24
;	PRIOR APPLICATION NUMBER: 60/090542
;	PRIOR FILING DATE: 1998-06-24
;	PRIOR APPLICATION NUMBER: 60/090557
;	PRIOR FILING DATE: 1998-06-24
;	PRIOR APPLICATION NUMBER: 60/090676
;	PRIOR FILING DATE: 1998-06-25
;	PRIOR APPLICATION NUMBER: 60/090678
;	PRIOR FILING DATE: 1998-06-25
;	PRIOR APPLICATION NUMBER: 60/090690
;	PRIOR FILING DATE: 1998-06-25
;	PRIOR APPLICATION NUMBER: 60/090694
;	PRIOR FILING DATE: 1998-06-25
;	PRIOR APPLICATION NUMBER: 60/090695

Sun Feb 16 09:13:26 2003

; PRIOR FILING DATE: 1998-06-25
; PRIOR APPLICATION NUMBER: 60/090696
; PRIOR FILING DATE: 1998-06-25
; PRIOR APPLICATION NUMBER: 60/090862
; PRIOR FILING DATE: 1998-06-26
; PRIOR APPLICATION NUMBER: 60/090863
; PRIOR FILING DATE: 1998-06-26
; PRIOR APPLICATION NUMBER: 60/091360
; PRIOR FILING DATE: 1998-07-01
; PRIOR APPLICATION NUMBER: 60/091478
; PRIOR FILING DATE: 1998-07-02
; PRIOR APPLICATION NUMBER: 60/091544
; PRIOR FILING DATE: 1998-07-01
; PRIOR APPLICATION NUMBER: 60/091519
; PRIOR FILING DATE: 1998-07-02
; PRIOR APPLICATION NUMBER: 60/091626
; PRIOR FILING DATE: 1998-07-02
; PRIOR APPLICATION NUMBER: 60/091633
; PRIOR FILING DATE: 1998-07-02
; PRIOR APPLICATION NUMBER: 60/091978
; PRIOR FILING DATE: 1998-07-07
; PRIOR APPLICATION NUMBER: 60/091982
; PRIOR FILING DATE: 1998-07-07
; PRIOR APPLICATION NUMBER: 60/092182
; PRIOR FILING DATE: 1998-07-09

Alignment Scores:
Pred. No.: 1.9e-05 Length: 762
Score: 112.00 Matches: 24
Percent Similarity: 53.57% Conservative: 6
Best Local Similarity: 42.86% Mismatches: 24
Query Match: 34.25% Indels: 2
DB: 9 Gaps: 2

09-833799-13B (1-57) x US-09-997-428-344 (1-762)

Qy 1 AlaGlnGluProValLysGlyProValSerThrLysProGlySerCysProIleLeu 20
Db 62 GCTGTGAAGGAGTTAAAGAGGGTATAGAG---AAAGCAGGGGTTTGCCCGAGCTGACAAC 118
Qy 21 IleArgCysAlaMetLeuAsnProProAsnArgCysLeuLysAspThrAspCysProGly 40
Db 119 GTACGCTGCTCAAGTCGATCCGATCTCC---CAGTGTCAACAGACCAGGACTGTCTGGGG 175
Qy 41 IleLysLysCysCysGluGlySerCysGlyMetAlaCysPheValPro 56
Db 176 GAAAGGAAGTGTGTATTACCTGCATGTGGCTTCAAGTGTGTGATTCCT 223

RESULT 22

US-09-997-666-344
; Sequence 344, Application US/09997666
; Publication No. US20030027163A1
; GENERAL INFORMATION:
; APPLICANT: Ashkenazi, Avi J.
; APPLICANT: Baker, Kevin P.
; APPLICANT: Botstein, David
; APPLICANT: Desnoyers, Luc
; APPLICANT: Eaton, Dan L.
; APPLICANT: Ferrara, Napoleone
; APPLICANT: Fong, Sherman
; APPLICANT: Gerber, Hanspeter
; APPLICANT: Gerritsen, Mary E.
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Grimaldi, J. Christopher
; APPLICANT: Gurney, Austin L.
; APPLICANT: Kljavin, Ivar J.
; APPLICANT: Napier, Mary A.
; APPLICANT: Pan, James
; APPLICANT: Paoni, Nicholas F.
; APPLICANT: Roy, Margaret Ann
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel

; APPLICANT: Watanabe, Colin K.
; APPLICANT: Williams, P. Mickey
; APPLICANT: Wood, William I.
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
; FILE REFERENCE: P2730P1C42
; CURRENT APPLICATION NUMBER: US/09/997,666
; PRIOR FILING DATE: 2001-11-15
; PRIOR APPLICATION NUMBER: 60/049787
; PRIOR FILING DATE: 1997-06-16
; PRIOR APPLICATION NUMBER: 60/062250
; PRIOR FILING DATE: 1997-10-17
; PRIOR APPLICATION NUMBER: 60/065186
; PRIOR FILING DATE: 1997-11-12
; PRIOR APPLICATION NUMBER: 60/065311
; PRIOR FILING DATE: 1997-11-13
; PRIOR APPLICATION NUMBER: 60/066770
; PRIOR FILING DATE: 1997-11-24
; PRIOR APPLICATION NUMBER: 60/075945
; PRIOR FILING DATE: 1998-02-25
; PRIOR APPLICATION NUMBER: 60/078910
; PRIOR FILING DATE: 1998-03-20
; PRIOR APPLICATION NUMBER: 60/083322
; PRIOR FILING DATE: 1998-04-28
; PRIOR APPLICATION NUMBER: 60/084600
; PRIOR FILING DATE: 1998-05-07
; PRIOR APPLICATION NUMBER: 60/087106
; PRIOR FILING DATE: 1998-05-28
; PRIOR APPLICATION NUMBER: 60/087607
; PRIOR FILING DATE: 1998-06-02
; PRIOR APPLICATION NUMBER: 60/087609
; PRIOR FILING DATE: 1998-06-02
; PRIOR APPLICATION NUMBER: 60/087759
; PRIOR FILING DATE: 1998-06-02
; PRIOR APPLICATION NUMBER: 60/087827
; PRIOR FILING DATE: 1998-06-03
; PRIOR APPLICATION NUMBER: 60/088021
; PRIOR FILING DATE: 1998-06-04
; PRIOR APPLICATION NUMBER: 60/088025
; PRIOR FILING DATE: 1998-06-04
; PRIOR APPLICATION NUMBER: 60/088026
; PRIOR FILING DATE: 1998-06-04
; PRIOR APPLICATION NUMBER: 60/088028
; PRIOR FILING DATE: 1998-06-04
; PRIOR APPLICATION NUMBER: 60/088029
; PRIOR FILING DATE: 1998-06-04
; PRIOR APPLICATION NUMBER: 60/088030
; PRIOR FILING DATE: 1998-06-04
; PRIOR APPLICATION NUMBER: 60/088033
; PRIOR FILING DATE: 1998-06-04
; PRIOR APPLICATION NUMBER: 60/088326
; PRIOR FILING DATE: 1998-06-04
; PRIOR APPLICATION NUMBER: 60/088167
; PRIOR FILING DATE: 1998-06-05
; PRIOR APPLICATION NUMBER: 60/088202
; PRIOR FILING DATE: 1998-06-05
; PRIOR APPLICATION NUMBER: 60/088212
; PRIOR FILING DATE: 1998-06-05
; PRIOR APPLICATION NUMBER: 60/088217
; PRIOR FILING DATE: 1998-06-05
; PRIOR APPLICATION NUMBER: 60/088655
; PRIOR FILING DATE: 1998-06-09
; PRIOR APPLICATION NUMBER: 60/088734
; PRIOR FILING DATE: 1998-06-10
; PRIOR APPLICATION NUMBER: 60/088738
; PRIOR FILING DATE: 1998-06-10
; PRIOR APPLICATION NUMBER: 60/088742
; PRIOR FILING DATE: 1998-06-10
; PRIOR APPLICATION NUMBER: 60/088810
; PRIOR FILING DATE: 1998-06-10
; PRIOR APPLICATION NUMBER: 60/088824
; PRIOR FILING DATE: 1998-06-10

; PRIOR APPLICATION NUMBER: 60/088826
; PRIOR FILING DATE: 1998-06-10
; PRIOR APPLICATION NUMBER: 60/088858
; PRIOR FILING DATE: 1998-06-11
; PRIOR APPLICATION NUMBER: 60/088861
; PRIOR FILING DATE: 1998-06-11
; PRIOR APPLICATION NUMBER: 60/088876
; PRIOR FILING DATE: 1998-06-11
; PRIOR APPLICATION NUMBER: 60/089105
; PRIOR FILING DATE: 1998-06-12
; PRIOR APPLICATION NUMBER: 60/089440
; PRIOR FILING DATE: 1998-06-16
; PRIOR APPLICATION NUMBER: 60/089512
; PRIOR FILING DATE: 1998-06-16
; PRIOR APPLICATION NUMBER: 60/089514
; PRIOR FILING DATE: 1998-06-16
; PRIOR APPLICATION NUMBER: 60/089532
; PRIOR FILING DATE: 1998-06-17
; PRIOR APPLICATION NUMBER: 60/089538
; PRIOR FILING DATE: 1998-06-17
; PRIOR APPLICATION NUMBER: 60/089598
; PRIOR FILING DATE: 1998-06-17
; PRIOR APPLICATION NUMBER: 60/089599
; PRIOR FILING DATE: 1998-06-17
; PRIOR APPLICATION NUMBER: 60/089600
; PRIOR FILING DATE: 1998-06-17
; PRIOR APPLICATION NUMBER: 60/089653
; PRIOR FILING DATE: 1998-06-17
; PRIOR APPLICATION NUMBER: 60/089801
; PRIOR FILING DATE: 1998-06-18
; PRIOR APPLICATION NUMBER: 60/089907
; PRIOR FILING DATE: 1998-06-18
; PRIOR APPLICATION NUMBER: 60/089908
; PRIOR FILING DATE: 1998-06-18
; PRIOR APPLICATION NUMBER: 60/089947
; PRIOR FILING DATE: 1998-06-19
; PRIOR APPLICATION NUMBER: 60/089948
; PRIOR FILING DATE: 1998-06-19
; PRIOR APPLICATION NUMBER: 60/089952
; PRIOR FILING DATE: 1998-06-19
; PRIOR APPLICATION NUMBER: 60/090246
; PRIOR FILING DATE: 1998-06-22
; PRIOR APPLICATION NUMBER: 60/090252
; PRIOR FILING DATE: 1998-06-22
; PRIOR APPLICATION NUMBER: 60/090254
; PRIOR FILING DATE: 1998-06-22
; PRIOR APPLICATION NUMBER: 60/090349
; PRIOR FILING DATE: 1998-06-23
; PRIOR APPLICATION NUMBER: 60/090355
; PRIOR FILING DATE: 1998-06-23
; PRIOR APPLICATION NUMBER: 60/090429
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090431
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090435
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090444
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090445
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090472
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090535
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090540
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090542
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090557
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090676
; PRIOR FILING DATE: 1998-06-25
; PRIOR APPLICATION NUMBER: 60/090678

; PRIOR FILING DATE: 1998-06-25
; PRIOR APPLICATION NUMBER: 60/090690
; PRIOR FILING DATE: 1998-06-25
; PRIOR APPLICATION NUMBER: 60/090694
; PRIOR FILING DATE: 1998-06-25
; PRIOR APPLICATION NUMBER: 60/090695
; PRIOR FILING DATE: 1998-06-25
; PRIOR APPLICATION NUMBER: 60/090696
; PRIOR FILING DATE: 1998-06-25
; PRIOR APPLICATION NUMBER: 60/090862
; PRIOR FILING DATE: 1998-06-26
; PRIOR APPLICATION NUMBER: 60/090863
; PRIOR FILING DATE: 1998-06-26
; PRIOR APPLICATION NUMBER: 60/091360
; PRIOR FILING DATE: 1998-07-01
; PRIOR APPLICATION NUMBER: 60/091478
; PRIOR FILING DATE: 1998-07-02
; PRIOR APPLICATION NUMBER: 60/091544
; PRIOR FILING DATE: 1998-07-01
; PRIOR APPLICATION NUMBER: 60/091519
; PRIOR FILING DATE: 1998-07-02
; PRIOR APPLICATION NUMBER: 60/091626
; PRIOR FILING DATE: 1998-07-02
; PRIOR APPLICATION NUMBER: 60/091633
; PRIOR FILING DATE: 1998-07-02
; PRIOR APPLICATION NUMBER: 60/091978
; PRIOR FILING DATE: 1998-07-07
; PRIOR APPLICATION NUMBER: 60/091982
; PRIOR FILING DATE: 1998-07-07
; PRIOR APPLICATION NUMBER: 60/092182
; PRIOR FILING DATE: 1998-07-09

Alignment Scores:

Pred. No.:	1.9e-05	Length:	762
Score:	112.00	Matches:	24
Percent Similarity:	53.57%	Conservative:	6
Best Local Similarity:	42.86%	Mismatches:	24
Query Match:	34.25%	Indels:	2
DB:	9	Gaps:	2

09-833799-13B (1-57) x US-09-997-666-344 (1-762)

Qy	1	AlaGlnGluProValLysGlyProValSerThrLysProGlySerCysProIleLeu	20
Db	62	GCTGTGAAGGAGTTAAAGAGGGGTATAGAG---AAAGCAGGGGTTTGGCCAGCTGACAAAC	118
Qy	21	IleArgCysAlaMetLeuAsnProAsnArgCysLeuLysAspThrAspCysProGly	40
Db	119	GTACGCTGCTTCAAGTCCGATCTCTCCC---CAGTGTACACAGACCAGGACTGTCTGGGG	175
Qy	41	IleLysLysCysCysGluGlySerCysGlyMetAlaCysPheValPro	56
Db	176	GAAAGGAAGTGTGTTACCTGCACACTGTGGCTTCAAGTGTGTGATTCTCT	223

RESULT 23

US-09-989-722-344
; Sequence 344, Application US/09989722
; Patent No. US20020072067A1
; GENERAL INFORMATION:
; APPLICANT: Ashkenazi, Avi J.
; APPLICANT: Baker, Kevin P.
; APPLICANT: Botstein, David
; APPLICANT: Desnoyers, Luc
; APPLICANT: Eaton, Dan L.
; APPLICANT: Ferrara, Napoleone
; APPLICANT: Fong, Sherman
; APPLICANT: Gerber, Hanspeter
; APPLICANT: Gerritsen, Mary E.
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Grimaldi, J. Christopher
; APPLICANT: Gurney, Austin L.
; APPLICANT: Kljavin, Ivar J.

APPLICANT: Napier, Mary A.
APPLICANT: Pan, James
APPLICANT: Paoni, Nicholas F.
APPLICANT: Roy, Margaret Ann
APPLICANT: Stewart, Timothy A.
APPLICANT: Tumas, Daniel
APPLICANT: Watanabe, Colin K.
APPLICANT: Williams, P. Mickey
APPLICANT: Wood, William I.
APPLICANT: Zhang, Zemin
TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
TITLE OF INVENTION: Acids Encoding the Same
FILE REFERENCE: P2730P1C63
CURRENT APPLICATION NUMBER: US/09/989,722
CURRENT FILING DATE: 2001-11-19
PRIOR APPLICATION NUMBER: 60/049787
PRIOR FILING DATE: 1997-06-16
PRIOR APPLICATION NUMBER: 60/062250
PRIOR FILING DATE: 1997-10-17
PRIOR APPLICATION NUMBER: 60/065186
PRIOR FILING DATE: 1997-11-12
PRIOR APPLICATION NUMBER: 60/065311
PRIOR FILING DATE: 1997-11-13
PRIOR APPLICATION NUMBER: 60/066770
PRIOR FILING DATE: 1997-11-24
PRIOR APPLICATION NUMBER: 60/075945
PRIOR FILING DATE: 1998-02-25
PRIOR APPLICATION NUMBER: 60/078910
PRIOR FILING DATE: 1998-03-20
PRIOR APPLICATION NUMBER: 60/083322
PRIOR FILING DATE: 1998-04-28
PRIOR APPLICATION NUMBER: 60/084600
PRIOR FILING DATE: 1998-05-07
PRIOR APPLICATION NUMBER: 60/087106
PRIOR FILING DATE: 1998-05-28
PRIOR APPLICATION NUMBER: 60/087607
PRIOR FILING DATE: 1998-06-02
PRIOR APPLICATION NUMBER: 60/087609
PRIOR FILING DATE: 1998-06-02
PRIOR APPLICATION NUMBER: 60/087759
PRIOR FILING DATE: 1998-06-02
PRIOR APPLICATION NUMBER: 60/087827
PRIOR FILING DATE: 1998-06-03
PRIOR APPLICATION NUMBER: 60/088021
PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088025
PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088026
PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088028
PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088029
PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088030
PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088033
PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088326
PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088167
PRIOR FILING DATE: 1998-06-05
PRIOR APPLICATION NUMBER: 60/088202
PRIOR FILING DATE: 1998-06-05
PRIOR APPLICATION NUMBER: 60/088212
PRIOR FILING DATE: 1998-06-05
PRIOR APPLICATION NUMBER: 60/088217
PRIOR FILING DATE: 1998-06-05
PRIOR APPLICATION NUMBER: 60/088655
PRIOR FILING DATE: 1998-06-09
PRIOR APPLICATION NUMBER: 60/088734
PRIOR FILING DATE: 1998-06-10
PRIOR APPLICATION NUMBER: 60/088738
PRIOR FILING DATE: 1998-06-10

PRIOR APPLICATION NUMBER: 60/088742
PRIOR FILING DATE: 1998-06-10
PRIOR APPLICATION NUMBER: 60/088810
PRIOR FILING DATE: 1998-06-10
PRIOR APPLICATION NUMBER: 60/088824
PRIOR FILING DATE: 1998-06-10
PRIOR APPLICATION NUMBER: 60/088826
PRIOR FILING DATE: 1998-06-10
PRIOR APPLICATION NUMBER: 60/088858
PRIOR FILING DATE: 1998-06-11
PRIOR APPLICATION NUMBER: 60/088861
PRIOR FILING DATE: 1998-06-11
PRIOR APPLICATION NUMBER: 60/088876
PRIOR FILING DATE: 1998-06-11
PRIOR APPLICATION NUMBER: 60/089105
PRIOR FILING DATE: 1998-06-12
PRIOR APPLICATION NUMBER: 60/089440
PRIOR FILING DATE: 1998-06-16
PRIOR APPLICATION NUMBER: 60/089512
PRIOR FILING DATE: 1998-06-16
PRIOR APPLICATION NUMBER: 60/089514
PRIOR FILING DATE: 1998-06-16
PRIOR APPLICATION NUMBER: 60/089532
PRIOR FILING DATE: 1998-06-17
PRIOR APPLICATION NUMBER: 60/089538
PRIOR FILING DATE: 1998-06-17
PRIOR APPLICATION NUMBER: 60/089598
PRIOR FILING DATE: 1998-06-17
PRIOR APPLICATION NUMBER: 60/089599
PRIOR FILING DATE: 1998-06-17
PRIOR APPLICATION NUMBER: 60/089600
PRIOR FILING DATE: 1998-06-17
PRIOR APPLICATION NUMBER: 60/089653
PRIOR FILING DATE: 1998-06-17
PRIOR APPLICATION NUMBER: 60/089801
PRIOR FILING DATE: 1998-06-18
PRIOR APPLICATION NUMBER: 60/089907
PRIOR FILING DATE: 1998-06-18
PRIOR APPLICATION NUMBER: 60/089908
PRIOR FILING DATE: 1998-06-18
PRIOR APPLICATION NUMBER: 60/089947
PRIOR FILING DATE: 1998-06-19
PRIOR APPLICATION NUMBER: 60/089948
PRIOR FILING DATE: 1998-06-19
PRIOR APPLICATION NUMBER: 60/089952
PRIOR FILING DATE: 1998-06-19
PRIOR APPLICATION NUMBER: 60/090246
PRIOR FILING DATE: 1998-06-22
PRIOR APPLICATION NUMBER: 60/090252
PRIOR FILING DATE: 1998-06-22
PRIOR APPLICATION NUMBER: 60/090254
PRIOR FILING DATE: 1998-06-22
PRIOR APPLICATION NUMBER: 60/090349
PRIOR FILING DATE: 1998-06-23
PRIOR APPLICATION NUMBER: 60/090355
PRIOR FILING DATE: 1998-06-23
PRIOR APPLICATION NUMBER: 60/090429
PRIOR FILING DATE: 1998-06-24
PRIOR APPLICATION NUMBER: 60/090431
PRIOR FILING DATE: 1998-06-24
PRIOR APPLICATION NUMBER: 60/090435
PRIOR FILING DATE: 1998-06-24
PRIOR APPLICATION NUMBER: 60/090444
PRIOR FILING DATE: 1998-06-24
PRIOR APPLICATION NUMBER: 60/090445
PRIOR FILING DATE: 1998-06-24
PRIOR APPLICATION NUMBER: 60/090472
PRIOR FILING DATE: 1998-06-24
PRIOR APPLICATION NUMBER: 60/090535
PRIOR FILING DATE: 1998-06-24
PRIOR APPLICATION NUMBER: 60/090540
PRIOR FILING DATE: 1998-06-24
PRIOR APPLICATION NUMBER: 60/090542

;
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090557
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090676
; PRIOR FILING DATE: 1998-06-25
; PRIOR APPLICATION NUMBER: 60/090678
; PRIOR FILING DATE: 1998-06-25
; PRIOR APPLICATION NUMBER: 60/090690
; PRIOR FILING DATE: 1998-06-25
; PRIOR APPLICATION NUMBER: 60/090694
; PRIOR FILING DATE: 1998-06-25
; PRIOR APPLICATION NUMBER: 60/090695
; PRIOR FILING DATE: 1998-06-25
; PRIOR APPLICATION NUMBER: 60/090696
; PRIOR FILING DATE: 1998-06-25
; PRIOR APPLICATION NUMBER: 60/090862
; PRIOR FILING DATE: 1998-06-26
; PRIOR APPLICATION NUMBER: 60/090863
; PRIOR FILING DATE: 1998-06-26
; PRIOR APPLICATION NUMBER: 60/091360
; PRIOR FILING DATE: 1998-07-01
; PRIOR APPLICATION NUMBER: 60/091478
; PRIOR FILING DATE: 1998-07-02
; PRIOR APPLICATION NUMBER: 60/091544
; PRIOR FILING DATE: 1998-07-01
; PRIOR APPLICATION NUMBER: 60/091519
; PRIOR FILING DATE: 1998-07-02
; PRIOR APPLICATION NUMBER: 60/091626
; PRIOR FILING DATE: 1998-07-02
; PRIOR APPLICATION NUMBER: 60/091633
; PRIOR FILING DATE: 1998-07-02
; PRIOR APPLICATION NUMBER: 60/091978
; PRIOR FILING DATE: 1998-07-07
; PRIOR APPLICATION NUMBER: 60/091982
; PRIOR FILING DATE: 1998-07-07
; PRIOR APPLICATION NUMBER: 60/092182
; PRIOR FILING DATE: 1998-07-09

Alignment Scores:

Pred. No.:	1.9e-05	Length:	762
Score:	112.00	Matches:	24
Percent Similarity:	53.57%	Conservative:	6
Best Local Similarity:	42.86%	Mismatches:	24
Query Match:	34.25%	Indels:	2
DB:	10	Gaps:	2

09-833799-13B (1-57) x US-09-989-722-344 (1-762)

QY 1 AlaGlnGluProValLysGlyProValSerThrLysProGlySerCysProIleLeu 20
Db 62 GCTGTGAAGGAGTTAAAGAGGGTATAGAG---AAAGCAGGGGTTGCCAGCTGACAC 118
QY 21 IleArgCysAlaMetLeuAsnProProAsnArgCysLeuLysAspThrAspCysProGly 40
Db 119 GTACGCTGCTTCAAGTCCGATCCTCCC---CAGTGTCAACACAGACCAGGACTGTCTGGG 175
QY 41 IleLysLysCysCysGluGlySerCysGlyMetAlaCysPheValPro 56
Db 176 GAAAGGAAGTGTGTACCTGCACTGTGGCTTCAAGTGTGTGATTCT 223

RESULT 24

US-09-989-723-344
; Sequence 344, Application US/09989723
; Patent No. US2002072092A1
; GENERAL INFORMATION:
; APPLICANT: Ashkenazi, Avi J.
; APPLICANT: Baker, Kevin P.
; APPLICANT: Botstein, David
; APPLICANT: Desnoyers, Luc
; APPLICANT: Eaton, Dan L.
; APPLICANT: Ferrara, Napoleone
; APPLICANT: Fong, Sherman
; APPLICANT: Gerber, Hanspeter

;
; APPLICANT: Gerritsen, Mary E.
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Grimaldi, J. Christopher
; APPLICANT: Gurney, Austin L.
; APPLICANT: Kljavin, Ivar J.
; APPLICANT: Napier, Mary A.
; APPLICANT: Pan, James
; APPLICANT: Paoni, Nicholas F.
; APPLICANT: Roy, Margaret Ann
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Watanabe, Colin K.
; APPLICANT: Williams, P. Mickey
; APPLICANT: Wood, William I.
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: Secrated and Transmembrane Polypeptides and Nucleic
; TITLE OF INVENTION: Acids Encoding the Same
; FILE REFERENCE: P2730P1C62
; CURRENT APPLICATION NUMBER: US/09/989,723
; CURRENT FILING DATE: 2001-11-19
; PRIOR APPLICATION NUMBER: 60/049787
; PRIOR FILING DATE: 1997-06-16
; PRIOR APPLICATION NUMBER: 60/062250
; PRIOR FILING DATE: 1997-10-17
; PRIOR APPLICATION NUMBER: 60/065186
; PRIOR FILING DATE: 1997-11-12
; PRIOR APPLICATION NUMBER: 60/065311
; PRIOR FILING DATE: 1997-11-13
; PRIOR APPLICATION NUMBER: 60/066770
; PRIOR FILING DATE: 1997-11-24
; PRIOR APPLICATION NUMBER: 60/075945
; PRIOR FILING DATE: 1998-02-25
; PRIOR APPLICATION NUMBER: 60/078910
; PRIOR FILING DATE: 1998-03-20
; PRIOR APPLICATION NUMBER: 60/083322
; PRIOR FILING DATE: 1998-04-28
; PRIOR APPLICATION NUMBER: 60/084600
; PRIOR FILING DATE: 1998-05-07
; PRIOR APPLICATION NUMBER: 60/087106
; PRIOR FILING DATE: 1998-05-28
; PRIOR APPLICATION NUMBER: 60/087607
; PRIOR FILING DATE: 1998-06-02
; PRIOR APPLICATION NUMBER: 60/087609
; PRIOR FILING DATE: 1998-06-02
; PRIOR APPLICATION NUMBER: 60/087759
; PRIOR FILING DATE: 1998-06-02
; PRIOR APPLICATION NUMBER: 60/087827
; PRIOR FILING DATE: 1998-06-03
; PRIOR APPLICATION NUMBER: 60/088021
; PRIOR FILING DATE: 1998-06-04
; PRIOR APPLICATION NUMBER: 60/088025
; PRIOR FILING DATE: 1998-06-04
; PRIOR APPLICATION NUMBER: 60/088026
; PRIOR FILING DATE: 1998-06-04
; PRIOR APPLICATION NUMBER: 60/088028
; PRIOR FILING DATE: 1998-06-04
; PRIOR APPLICATION NUMBER: 60/088029
; PRIOR FILING DATE: 1998-06-04
; PRIOR APPLICATION NUMBER: 60/088030
; PRIOR FILING DATE: 1998-06-04
; PRIOR APPLICATION NUMBER: 60/088033
; PRIOR FILING DATE: 1998-06-04
; PRIOR APPLICATION NUMBER: 60/088326
; PRIOR FILING DATE: 1998-06-04
; PRIOR APPLICATION NUMBER: 60/088167
; PRIOR FILING DATE: 1998-06-05
; PRIOR APPLICATION NUMBER: 60/088202
; PRIOR FILING DATE: 1998-06-05
; PRIOR APPLICATION NUMBER: 60/088212
; PRIOR FILING DATE: 1998-06-05
; PRIOR APPLICATION NUMBER: 60/088217
; PRIOR FILING DATE: 1998-06-05

Sun Feb 16 09:13:26 2003

;
; PRIOR APPLICATION NUMBER: 60/088655
; PRIOR FILING DATE: 1998-06-09
; PRIOR APPLICATION NUMBER: 60/088734
; PRIOR FILING DATE: 1998-06-10
; PRIOR APPLICATION NUMBER: 60/088738
; PRIOR FILING DATE: 1998-06-10
; PRIOR APPLICATION NUMBER: 60/088742
; PRIOR FILING DATE: 1998-06-10
; PRIOR APPLICATION NUMBER: 60/088810
; PRIOR FILING DATE: 1998-06-10
; PRIOR APPLICATION NUMBER: 60/088824
; PRIOR FILING DATE: 1998-06-10
; PRIOR APPLICATION NUMBER: 60/088826
; PRIOR FILING DATE: 1998-06-10
; PRIOR APPLICATION NUMBER: 60/088858
; PRIOR FILING DATE: 1998-06-11
; PRIOR APPLICATION NUMBER: 60/088861
; PRIOR FILING DATE: 1998-06-11
; PRIOR APPLICATION NUMBER: 60/088876
; PRIOR FILING DATE: 1998-06-11
; PRIOR APPLICATION NUMBER: 60/089105
; PRIOR FILING DATE: 1998-06-12
; PRIOR APPLICATION NUMBER: 60/089440
; PRIOR FILING DATE: 1998-06-16
; PRIOR APPLICATION NUMBER: 60/089512
; PRIOR FILING DATE: 1998-06-16
; PRIOR APPLICATION NUMBER: 60/089514
; PRIOR FILING DATE: 1998-06-16
; PRIOR APPLICATION NUMBER: 60/089532
; PRIOR FILING DATE: 1998-06-17
; PRIOR APPLICATION NUMBER: 60/089538
; PRIOR FILING DATE: 1998-06-17
; PRIOR APPLICATION NUMBER: 60/089598
; PRIOR FILING DATE: 1998-06-17
; PRIOR APPLICATION NUMBER: 60/089599
; PRIOR FILING DATE: 1998-06-17
; PRIOR APPLICATION NUMBER: 60/089600
; PRIOR FILING DATE: 1998-06-17
; PRIOR APPLICATION NUMBER: 60/089653
; PRIOR FILING DATE: 1998-06-17
; PRIOR APPLICATION NUMBER: 60/089801
; PRIOR FILING DATE: 1998-06-18
; PRIOR APPLICATION NUMBER: 60/089907
; PRIOR FILING DATE: 1998-06-18
; PRIOR APPLICATION NUMBER: 60/089908
; PRIOR FILING DATE: 1998-06-18
; PRIOR APPLICATION NUMBER: 60/089947
; PRIOR FILING DATE: 1998-06-19
; PRIOR APPLICATION NUMBER: 60/089948
; PRIOR FILING DATE: 1998-06-19
; PRIOR APPLICATION NUMBER: 60/089952
; PRIOR FILING DATE: 1998-06-19
; PRIOR APPLICATION NUMBER: 60/090246
; PRIOR FILING DATE: 1998-06-22
; PRIOR APPLICATION NUMBER: 60/090252
; PRIOR FILING DATE: 1998-06-22
; PRIOR APPLICATION NUMBER: 60/090254
; PRIOR FILING DATE: 1998-06-22
; PRIOR APPLICATION NUMBER: 60/090349
; PRIOR FILING DATE: 1998-06-23
; PRIOR APPLICATION NUMBER: 60/090355
; PRIOR FILING DATE: 1998-06-23
; PRIOR APPLICATION NUMBER: 60/090429
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090431
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090435
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090444
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090445
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090472

;
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090535
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090540
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090542
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090557
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090676
; PRIOR FILING DATE: 1998-06-25
; PRIOR APPLICATION NUMBER: 60/090678
; PRIOR FILING DATE: 1998-06-25
; PRIOR APPLICATION NUMBER: 60/090690
; PRIOR FILING DATE: 1998-06-25
; PRIOR APPLICATION NUMBER: 60/090694
; PRIOR FILING DATE: 1998-06-25
; PRIOR APPLICATION NUMBER: 60/090695
; PRIOR FILING DATE: 1998-06-25
; PRIOR APPLICATION NUMBER: 60/090696
; PRIOR FILING DATE: 1998-06-25
; PRIOR APPLICATION NUMBER: 60/090862
; PRIOR FILING DATE: 1998-06-26
; PRIOR APPLICATION NUMBER: 60/090863
; PRIOR FILING DATE: 1998-06-26
; PRIOR APPLICATION NUMBER: 60/091360
; PRIOR FILING DATE: 1998-07-01
; PRIOR APPLICATION NUMBER: 60/091478
; PRIOR FILING DATE: 1998-07-02
; PRIOR APPLICATION NUMBER: 60/091544
; PRIOR FILING DATE: 1998-07-01
; PRIOR APPLICATION NUMBER: 60/091519
; PRIOR FILING DATE: 1998-07-02
; PRIOR APPLICATION NUMBER: 60/091626
; PRIOR FILING DATE: 1998-07-02
; PRIOR APPLICATION NUMBER: 60/091633
; PRIOR FILING DATE: 1998-07-02
; PRIOR APPLICATION NUMBER: 60/091978
; PRIOR FILING DATE: 1998-07-07
; PRIOR APPLICATION NUMBER: 60/091982
; PRIOR FILING DATE: 1998-07-07
; PRIOR APPLICATION NUMBER: 60/092182
; PRIOR FILING DATE: 1998-07-09

Alignment Scores:

Pred. No.: 1.9e-05 Length: 762
Score: 112.00 Matches: 24
Percent Similarity: 53.57% Conservative: 6
Best Local Similarity: 42.86% Mismatches: 24
Query Match: 34.25% Indels: 2
DB: 10 Gaps: 2

09-833799-13B (1-57) x US-09-989-723-344 (1-762)

QY 1 AlaGlnGluProValLysGlyProValSerThrLysProGlySerCysProIleLeu 20
Db 62 GCTGTGGAAGGAGTTAAAGAGGGTATAGAG--AAAGCAGGGGTTGCCAGCTGACAAC 118
QY 21 IleArgCysAlaMetLeuAsnProProAsnArgCysLeuLysAspThrAspCysProGly 40
Db 119 GTACGCTGCTTCAAGTCCGATCCTCCC---CAGTGTACACAGACCAGGACTGTCTGGGG 175
QY 41 IleLysLysCysCysGluGlySerCysGlyMetAlaCysPheValPro 56
Db 176 GAAAGGAAGTGTGTACCTGCACTGTGGCTTCAAGTGTGTGATTCCT 223

RESULT 25

US-09-989-279-344
; Sequence 344, Application US/09989279
; Patent No. US20020072496A1
; GENERAL INFORMATION:
; APPLICANT: Ashkenazi, Avi J.
; APPLICANT: Baker, Kevin P.

; APPLICANT: Botstein, David
; APPLICANT: Desnoyers, Luc
; APPLICANT: Eaton, Dan L.
; APPLICANT: Ferrara, Napoleone
; APPLICANT: Fong, Sherman
; APPLICANT: Gerber, Hanspeter
; APPLICANT: Gerritsen, Mary E.
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Grimaldi, J. Christopher
; APPLICANT: Gurney, Austin L.
; APPLICANT: Kljavin, Ivar J.
; APPLICANT: Napier, Mary A.
; APPLICANT: Pan, James
; APPLICANT: Paoni, Nicholas F.
; APPLICANT: Roy, Margaret Ann
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Watanabe, Colin K.
; APPLICANT: Williams, P. Mickey
; APPLICANT: Wood, William I.
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
; TITLE OF INVENTION: Acids Encoding the Same
; FILE REFERENCE: P2730P1C56
; CURRENT APPLICATION NUMBER: US/09/989,279
; CURRENT FILING DATE: 2001-11-19
; PRIOR APPLICATION NUMBER: 60/049787
; PRIOR FILING DATE: 1997-06-16
; PRIOR APPLICATION NUMBER: 60/062250
; PRIOR FILING DATE: 1997-10-17
; PRIOR APPLICATION NUMBER: 60/065186
; PRIOR FILING DATE: 1997-11-12
; PRIOR APPLICATION NUMBER: 60/065311
; PRIOR FILING DATE: 1997-11-13
; PRIOR APPLICATION NUMBER: 60/066770
; PRIOR FILING DATE: 1997-11-24
; PRIOR APPLICATION NUMBER: 60/075945
; PRIOR FILING DATE: 1998-02-25
; PRIOR APPLICATION NUMBER: 60/078910
; PRIOR FILING DATE: 1998-03-20
; PRIOR APPLICATION NUMBER: 60/083322
; PRIOR FILING DATE: 1998-04-28
; PRIOR APPLICATION NUMBER: 60/084600
; PRIOR FILING DATE: 1998-05-07
; PRIOR APPLICATION NUMBER: 60/087106
; PRIOR FILING DATE: 1998-05-28
; PRIOR APPLICATION NUMBER: 60/087607
; PRIOR FILING DATE: 1998-06-02
; PRIOR APPLICATION NUMBER: 60/087609
; PRIOR FILING DATE: 1998-06-02
; PRIOR APPLICATION NUMBER: 60/087759
; PRIOR FILING DATE: 1998-06-02
; PRIOR APPLICATION NUMBER: 60/087827
; PRIOR FILING DATE: 1998-06-03
; PRIOR APPLICATION NUMBER: 60/088021
; PRIOR FILING DATE: 1998-06-04
; PRIOR APPLICATION NUMBER: 60/088025
; PRIOR FILING DATE: 1998-06-04
; PRIOR APPLICATION NUMBER: 60/088026
; PRIOR FILING DATE: 1998-06-04
; PRIOR APPLICATION NUMBER: 60/088028
; PRIOR FILING DATE: 1998-06-04
; PRIOR APPLICATION NUMBER: 60/088029
; PRIOR FILING DATE: 1998-06-04
; PRIOR APPLICATION NUMBER: 60/088030
; PRIOR FILING DATE: 1998-06-04
; PRIOR APPLICATION NUMBER: 60/088033
; PRIOR FILING DATE: 1998-06-04
; PRIOR APPLICATION NUMBER: 60/088326
; PRIOR FILING DATE: 1998-06-04
; PRIOR APPLICATION NUMBER: 60/088167
; PRIOR FILING DATE: 1998-06-05
; PRIOR APPLICATION NUMBER: 60/088202
; PRIOR FILING DATE: 1998-06-05
; PRIOR APPLICATION NUMBER: 60/088212
; PRIOR FILING DATE: 1998-06-05
; PRIOR APPLICATION NUMBER: 60/088217
; PRIOR FILING DATE: 1998-06-05
; PRIOR APPLICATION NUMBER: 60/088655
; PRIOR FILING DATE: 1998-06-09
; PRIOR APPLICATION NUMBER: 60/088734
; PRIOR FILING DATE: 1998-06-10
; PRIOR APPLICATION NUMBER: 60/088738
; PRIOR FILING DATE: 1998-06-10
; PRIOR APPLICATION NUMBER: 60/088742
; PRIOR FILING DATE: 1998-06-10
; PRIOR APPLICATION NUMBER: 60/088810
; PRIOR FILING DATE: 1998-06-10
; PRIOR APPLICATION NUMBER: 60/088824
; PRIOR FILING DATE: 1998-06-10
; PRIOR APPLICATION NUMBER: 60/088826
; PRIOR FILING DATE: 1998-06-10
; PRIOR APPLICATION NUMBER: 60/088858
; PRIOR FILING DATE: 1998-06-11
; PRIOR APPLICATION NUMBER: 60/088861
; PRIOR FILING DATE: 1998-06-11
; PRIOR APPLICATION NUMBER: 60/088876
; PRIOR FILING DATE: 1998-06-11
; PRIOR APPLICATION NUMBER: 60/089105
; PRIOR FILING DATE: 1998-06-12
; PRIOR APPLICATION NUMBER: 60/089440
; PRIOR FILING DATE: 1998-06-16
; PRIOR APPLICATION NUMBER: 60/089512
; PRIOR FILING DATE: 1998-06-16
; PRIOR APPLICATION NUMBER: 60/089514
; PRIOR FILING DATE: 1998-06-16
; PRIOR APPLICATION NUMBER: 60/089532
; PRIOR FILING DATE: 1998-06-17
; PRIOR APPLICATION NUMBER: 60/089538
; PRIOR FILING DATE: 1998-06-17
; PRIOR APPLICATION NUMBER: 60/089598
; PRIOR FILING DATE: 1998-06-17
; PRIOR APPLICATION NUMBER: 60/089599
; PRIOR FILING DATE: 1998-06-17
; PRIOR APPLICATION NUMBER: 60/089600
; PRIOR FILING DATE: 1998-06-17
; PRIOR APPLICATION NUMBER: 60/089653
; PRIOR FILING DATE: 1998-06-17
; PRIOR APPLICATION NUMBER: 60/089801
; PRIOR FILING DATE: 1998-06-18
; PRIOR APPLICATION NUMBER: 60/089907
; PRIOR FILING DATE: 1998-06-18
; PRIOR APPLICATION NUMBER: 60/089908
; PRIOR FILING DATE: 1998-06-18
; PRIOR APPLICATION NUMBER: 60/089947
; PRIOR FILING DATE: 1998-06-19
; PRIOR APPLICATION NUMBER: 60/089948
; PRIOR FILING DATE: 1998-06-19
; PRIOR APPLICATION NUMBER: 60/089952
; PRIOR FILING DATE: 1998-06-19
; PRIOR APPLICATION NUMBER: 60/090246
; PRIOR FILING DATE: 1998-06-22
; PRIOR APPLICATION NUMBER: 60/090252
; PRIOR FILING DATE: 1998-06-22
; PRIOR APPLICATION NUMBER: 60/090254
; PRIOR FILING DATE: 1998-06-22
; PRIOR APPLICATION NUMBER: 60/090349
; PRIOR FILING DATE: 1998-06-23
; PRIOR APPLICATION NUMBER: 60/090355
; PRIOR FILING DATE: 1998-06-23
; PRIOR APPLICATION NUMBER: 60/090429
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090431
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090435

; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090444
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090445
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090472
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090535
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090540
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090542
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090557
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090676
; PRIOR FILING DATE: 1998-06-25
; PRIOR APPLICATION NUMBER: 60/090678
; PRIOR FILING DATE: 1998-06-25
; PRIOR APPLICATION NUMBER: 60/090690
; PRIOR FILING DATE: 1998-06-25
; PRIOR APPLICATION NUMBER: 60/090694
; PRIOR FILING DATE: 1998-06-25
; PRIOR APPLICATION NUMBER: 60/090695
; PRIOR FILING DATE: 1998-06-25
; PRIOR APPLICATION NUMBER: 60/090696
; PRIOR FILING DATE: 1998-06-25
; PRIOR APPLICATION NUMBER: 60/090862
; PRIOR FILING DATE: 1998-06-26
; PRIOR APPLICATION NUMBER: 60/090863
; PRIOR FILING DATE: 1998-06-26
; PRIOR APPLICATION NUMBER: 60/091360
; PRIOR FILING DATE: 1998-07-01
; PRIOR APPLICATION NUMBER: 60/091478
; PRIOR FILING DATE: 1998-07-02
; PRIOR APPLICATION NUMBER: 60/091544
; PRIOR FILING DATE: 1998-07-01
; PRIOR APPLICATION NUMBER: 60/091519
; PRIOR FILING DATE: 1998-07-02
; PRIOR APPLICATION NUMBER: 60/091626
; PRIOR FILING DATE: 1998-07-02
; PRIOR APPLICATION NUMBER: 60/091633
; PRIOR FILING DATE: 1998-07-02
; PRIOR APPLICATION NUMBER: 60/091978
; PRIOR FILING DATE: 1998-07-07
; PRIOR APPLICATION NUMBER: 60/091982
; PRIOR FILING DATE: 1998-07-07
; PRIOR APPLICATION NUMBER: 60/092182
; PRIOR FILING DATE: 1998-07-09

Alignment Scores:

Pred. No.:	1.9e-05	Length:	762
Score:	112.00	Matches:	24
Percent Similarity:	53.57%	Conservative:	6
Best Local Similarity:	42.86%	Mismatches:	24
Query Match:	34.25%	Indels:	2
DB:	10	Gaps:	2

09-833799-13B (1-57) x US-09-989-279-344 (1-762)

QY	1	AlaGlnGluProValLysGlyProValSerThrLysProGlySerCysProIleLeu	20
Db	62	GCTGTGAAGGAGGTAAAGAGGGTATACAG---AAAGCAGGGGTTGCCAGCTGACAAC	118
QY	21	IleArgCysAlaMetLeuAsnProProAsnArgCysLeuLysAspThrAspCysProGly	40
Db	119	GTACGCTGCTCAAGTCCGATCCTCC---CAGTGTACACACAGCAGGACTGTCTGGG	175
QY	41	IleLysLysCysCysGluGlySerGlyMetAlaCysPheValPro	56
Db	176	GAAGGAAGTGTGTACCTGCCTGCTGCTTCAAGTGTGTGATTCCT	223

RESULT 26

US-09-989-727-344
; Sequence 344, Application US/09989727
; Patent No. US20020072497A1
; GENERAL INFORMATION:
; APPLICANT: Ashkenazi, Avi J.
; APPLICANT: Baker, Kevin P.
; APPLICANT: Botstein, David
; APPLICANT: Desnoyers, Luc
; APPLICANT: Eaton, Dan L.
; APPLICANT: Ferrara, Napoleone
; APPLICANT: Fong, Sherman
; APPLICANT: Gerber, Hanspeter
; APPLICANT: Gerritsen, Mary E.
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Grimaldi, J. Christopher
; APPLICANT: Gurney, Austin L.
; APPLICANT: Kljavin, Ivar J.
; APPLICANT: Napier, Mary A.
; APPLICANT: Pan, James
; APPLICANT: Paoni, Nicholas F.
; APPLICANT: Roy, Margaret Ann
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Watanabe, Colin K.
; APPLICANT: Williams, P. Mickey
; APPLICANT: Wood, William I.
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
; FILE REFERENCE: P2730P1C65
; CURRENT APPLICATION NUMBER: US/09/989,727
; CURRENT FILING DATE: 2001-11-19
; PRIOR APPLICATION NUMBER: 60/049787
; PRIOR FILING DATE: 1997-06-16
; PRIOR APPLICATION NUMBER: 60/062250
; PRIOR FILING DATE: 1997-10-17
; PRIOR APPLICATION NUMBER: 60/065186
; PRIOR FILING DATE: 1997-11-12
; PRIOR APPLICATION NUMBER: 60/065311
; PRIOR FILING DATE: 1997-11-13
; PRIOR APPLICATION NUMBER: 60/066770
; PRIOR FILING DATE: 1997-11-24
; PRIOR APPLICATION NUMBER: 60/075945
; PRIOR FILING DATE: 1998-02-25
; PRIOR APPLICATION NUMBER: 60/078910
; PRIOR FILING DATE: 1998-03-20
; PRIOR APPLICATION NUMBER: 60/083322
; PRIOR FILING DATE: 1998-04-28
; PRIOR APPLICATION NUMBER: 60/084600
; PRIOR FILING DATE: 1998-05-07
; PRIOR APPLICATION NUMBER: 60/087106
; PRIOR FILING DATE: 1998-05-28
; PRIOR APPLICATION NUMBER: 60/087607
; PRIOR FILING DATE: 1998-06-02
; PRIOR APPLICATION NUMBER: 60/087609
; PRIOR FILING DATE: 1998-06-02
; PRIOR APPLICATION NUMBER: 60/087759
; PRIOR FILING DATE: 1998-06-02
; PRIOR APPLICATION NUMBER: 60/087827
; PRIOR FILING DATE: 1998-06-03
; PRIOR APPLICATION NUMBER: 60/088021
; PRIOR FILING DATE: 1998-06-04
; PRIOR APPLICATION NUMBER: 60/088025
; PRIOR FILING DATE: 1998-06-04
; PRIOR APPLICATION NUMBER: 60/088026
; PRIOR FILING DATE: 1998-06-04
; PRIOR APPLICATION NUMBER: 60/088028
; PRIOR FILING DATE: 1998-06-04
; PRIOR APPLICATION NUMBER: 60/088029
; PRIOR FILING DATE: 1998-06-04
; PRIOR APPLICATION NUMBER: 60/088030
; PRIOR FILING DATE: 1998-06-04

QY 41 IleIysLysCysCysGluGlySerCysGlyMetAlaCysPheValPro 56
Db 176 GAAAGGAAGTGTGTTACCTGCACTGGCTTCAAGTGTGTGATTCCT 223

RESULT 27

US-09-989-731-344

; Sequence 344, Application US/09989731

; Patent No. US20020103125A1

; GENERAL INFORMATION:

; APPLICANT: Ashkenazi, Avi J.

; APPLICANT: Baker, Kevin P.

; APPLICANT: Botstein, David

; APPLICANT: Desnoyers, Luc

; APPLICANT: Baton, Dan L.

; APPLICANT: Ferrara, Napoleone

; APPLICANT: Fong, Sherman

; APPLICANT: Gerber, Hanspeter

; APPLICANT: Gerritsen, Mary E.

; APPLICANT: Goddard, Audrey

; APPLICANT: Godowski, Paul J.

; APPLICANT: Grimaldi, J. Christopher

; APPLICANT: Gurney, Austin L.

; APPLICANT: Kljavin, Ivar J.

; APPLICANT: Napier, Mary A.

; APPLICANT: Pan, James

; APPLICANT: Paoni, Nicholas F.

; APPLICANT: Roy, Margaret Ann

; APPLICANT: Stewart, Timothy A.

; APPLICANT: Tumas, Daniel

; APPLICANT: Watanabe, Colin K.

; APPLICANT: Williams, P. Mickey

; APPLICANT: Wood, William I.

; APPLICANT: Zhang, Zemin

; TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic

; TITLE OF INVENTION: Acids Encoding the Same

; FILE REFERENCE: P2730P1C70

; CURRENT APPLICATION NUMBER: US/09/989,731

; CURRENT FILING DATE: 2001-11-20

; PRIOR APPLICATION NUMBER: 60/049787

; PRIOR FILING DATE: 1997-06-16

; PRIOR APPLICATION NUMBER: 60/062250

; PRIOR FILING DATE: 1997-10-17

; PRIOR APPLICATION NUMBER: 60/065186

; PRIOR FILING DATE: 1997-11-12

; PRIOR APPLICATION NUMBER: 60/065311

; PRIOR FILING DATE: 1997-11-13

; PRIOR APPLICATION NUMBER: 60/066770

; PRIOR FILING DATE: 1997-11-24

; PRIOR APPLICATION NUMBER: 60/075945

; PRIOR FILING DATE: 1998-02-25

; PRIOR APPLICATION NUMBER: 60/078910

; PRIOR FILING DATE: 1998-03-20

; PRIOR APPLICATION NUMBER: 60/083322

; PRIOR FILING DATE: 1998-04-28

; PRIOR APPLICATION NUMBER: 60/084600

; PRIOR FILING DATE: 1998-05-07

; PRIOR APPLICATION NUMBER: 60/087106

; PRIOR FILING DATE: 1998-05-28

; PRIOR APPLICATION NUMBER: 60/087607

; PRIOR FILING DATE: 1998-06-02

; PRIOR APPLICATION NUMBER: 60/087609

; PRIOR FILING DATE: 1998-06-02

; PRIOR APPLICATION NUMBER: 60/087759

; PRIOR FILING DATE: 1998-06-02

; PRIOR APPLICATION NUMBER: 60/087827

; PRIOR FILING DATE: 1998-06-03

; PRIOR APPLICATION NUMBER: 60/088021

; PRIOR FILING DATE: 1998-06-04

; PRIOR APPLICATION NUMBER: 60/088025

; PRIOR FILING DATE: 1998-06-04

; PRIOR APPLICATION NUMBER: 60/088026

; PRIOR FILING DATE: 1998-06-04

; PRIOR APPLICATION NUMBER: 60/088028

; PRIOR FILING DATE: 1998-06-04

; PRIOR APPLICATION NUMBER: 60/088029

; PRIOR FILING DATE: 1998-06-04

; PRIOR APPLICATION NUMBER: 60/088030

; PRIOR FILING DATE: 1998-06-04

; PRIOR APPLICATION NUMBER: 60/088033

; PRIOR FILING DATE: 1998-06-04

; PRIOR APPLICATION NUMBER: 60/088326

; PRIOR FILING DATE: 1998-06-04

; PRIOR APPLICATION NUMBER: 60/088167

; PRIOR FILING DATE: 1998-06-05

; PRIOR APPLICATION NUMBER: 60/088202

; PRIOR FILING DATE: 1998-06-05

; PRIOR APPLICATION NUMBER: 60/088212

; PRIOR FILING DATE: 1998-06-05

; PRIOR APPLICATION NUMBER: 60/088217

; PRIOR FILING DATE: 1998-06-05

; PRIOR APPLICATION NUMBER: 60/088655

; PRIOR FILING DATE: 1998-06-09

; PRIOR APPLICATION NUMBER: 60/088734

; PRIOR FILING DATE: 1998-06-10

; PRIOR APPLICATION NUMBER: 60/088738

; PRIOR FILING DATE: 1998-06-10

; PRIOR APPLICATION NUMBER: 60/088742

; PRIOR FILING DATE: 1998-06-10

; PRIOR APPLICATION NUMBER: 60/088810

; PRIOR FILING DATE: 1998-06-10

; PRIOR APPLICATION NUMBER: 60/088824

; PRIOR FILING DATE: 1998-06-10

; PRIOR APPLICATION NUMBER: 60/088826

; PRIOR FILING DATE: 1998-06-10

; PRIOR APPLICATION NUMBER: 60/088858

; PRIOR FILING DATE: 1998-06-11

; PRIOR APPLICATION NUMBER: 60/088861

; PRIOR FILING DATE: 1998-06-11

; PRIOR APPLICATION NUMBER: 60/088876

; PRIOR FILING DATE: 1998-06-11

; PRIOR APPLICATION NUMBER: 60/089105

; PRIOR FILING DATE: 1998-06-12

; PRIOR APPLICATION NUMBER: 60/089440

; PRIOR FILING DATE: 1998-06-16

; PRIOR APPLICATION NUMBER: 60/089512

; PRIOR FILING DATE: 1998-06-16

; PRIOR APPLICATION NUMBER: 60/089514

; PRIOR FILING DATE: 1998-06-16

; PRIOR APPLICATION NUMBER: 60/089532

; PRIOR FILING DATE: 1998-06-17

; PRIOR APPLICATION NUMBER: 60/089538

; PRIOR FILING DATE: 1998-06-17

; PRIOR APPLICATION NUMBER: 60/089598

; PRIOR FILING DATE: 1998-06-17

; PRIOR APPLICATION NUMBER: 60/089599

; PRIOR FILING DATE: 1998-06-17

; PRIOR APPLICATION NUMBER: 60/089600

; PRIOR FILING DATE: 1998-06-17

; PRIOR APPLICATION NUMBER: 60/089653

; PRIOR FILING DATE: 1998-06-17

; PRIOR APPLICATION NUMBER: 60/089801

; PRIOR FILING DATE: 1998-06-18

; PRIOR APPLICATION NUMBER: 60/089907

; PRIOR FILING DATE: 1998-06-18

; PRIOR APPLICATION NUMBER: 60/089908

; PRIOR FILING DATE: 1998-06-18

; PRIOR APPLICATION NUMBER: 60/089947

; PRIOR FILING DATE: 1998-06-19

; PRIOR APPLICATION NUMBER: 60/089948

; PRIOR FILING DATE: 1998-06-19

; PRIOR APPLICATION NUMBER: 60/089952

; PRIOR FILING DATE: 1998-06-19

; PRIOR APPLICATION NUMBER: 60/090246

; PRIOR FILING DATE: 1998-06-22

; PRIOR APPLICATION NUMBER: 60/090252

;
; PRIOR FILING DATE: 1998-06-22
; PRIOR APPLICATION NUMBER: 60/090254
; PRIOR FILING DATE: 1998-06-22
; PRIOR APPLICATION NUMBER: 60/090349
; PRIOR FILING DATE: 1998-06-23
; PRIOR APPLICATION NUMBER: 60/090355
; PRIOR FILING DATE: 1998-06-23
; PRIOR APPLICATION NUMBER: 60/090429
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090431
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090435
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090444
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090445
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090472
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090535
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090540
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090542
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090557
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090676
; PRIOR FILING DATE: 1998-06-25
; PRIOR APPLICATION NUMBER: 60/090678
; PRIOR FILING DATE: 1998-06-25
; PRIOR APPLICATION NUMBER: 60/090690
; PRIOR FILING DATE: 1998-06-25
; PRIOR APPLICATION NUMBER: 60/090694
; PRIOR FILING DATE: 1998-06-25
; PRIOR APPLICATION NUMBER: 60/090695
; PRIOR FILING DATE: 1998-06-25
; PRIOR APPLICATION NUMBER: 60/090696
; PRIOR FILING DATE: 1998-06-25
; PRIOR APPLICATION NUMBER: 60/090862
; PRIOR FILING DATE: 1998-06-26
; PRIOR APPLICATION NUMBER: 60/090863
; PRIOR FILING DATE: 1998-06-26
; PRIOR APPLICATION NUMBER: 60/091360
; PRIOR FILING DATE: 1998-07-01
; PRIOR APPLICATION NUMBER: 60/091478
; PRIOR FILING DATE: 1998-07-02
; PRIOR APPLICATION NUMBER: 60/091544
; PRIOR FILING DATE: 1998-07-01
; PRIOR APPLICATION NUMBER: 60/091519
; PRIOR FILING DATE: 1998-07-02
; PRIOR APPLICATION NUMBER: 60/091626
; PRIOR FILING DATE: 1998-07-02
; PRIOR APPLICATION NUMBER: 60/091633
; PRIOR FILING DATE: 1998-07-02
; PRIOR APPLICATION NUMBER: 60/091978
; PRIOR FILING DATE: 1998-07-07
; PRIOR APPLICATION NUMBER: 60/091982
; PRIOR FILING DATE: 1998-07-07
; PRIOR APPLICATION NUMBER: 60/092182
; PRIOR FILING DATE: 1998-07-09

Alignment Scores:
Pred. No.: 1.9e-05 Length: 762
Score: 112.00 Matches: 24
Percent Similarity: 53.57% Conservative: 6
Best Local Similarity: 42.86% Mismatches: 24
Query Match: 34.25% Indels: 2
DB: 10 Gaps: 2

09-833799-13B (1-57) x US-09-989-731-344 (1-762)

QY 1 AlaGlnGluProValLysGlyProValSerThrLysProGlySerCysProIleLeu 20

Db 62 GCTGTGGAAGGAGTTAAAGAGGGGTATAGAG---AAAGCAGGGGTTTGCCAGCTGACAAC 118
QY 21 IleArgCysAlaMetLeuAsnProProAsnArgCysLeuLysAspThrAspCysProGly 40
Db 119 GTACGCTGCTTCAAGTCCGATCCTCCC---CAGTGTACACAGACCAGGACTGTCTGGGG 175
QY 41 IleLysLysCysCysGluGlySerCysGlyMetAlaCysPheValPro 56
Db 176 GAAAGGAAGTGTGTGTACCTGCACCTGTGGCTTCAAGTGTGTGATTCCT 223

RESULT 28

US-09-989-732-344
; Sequence 344, Application US/09989732
; Patent No. US20020123463A1

; GENERAL INFORMATION:

; APPLICANT: Ashkenazi, Avi J.
; APPLICANT: Baker, Kevin P.
; APPLICANT: Botstein, David
; APPLICANT: Desnoyers, Luc
; APPLICANT: Baton, Dan L.
; APPLICANT: Ferrara, Napoleone
; APPLICANT: Fong, Sherman
; APPLICANT: Gerber, Hanspeter
; APPLICANT: Gerritsen, Mary E.
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Grimaldi, J. Christopher
; APPLICANT: Gurney, Austin L.
; APPLICANT: Kljavin, Ivar J.
; APPLICANT: Napier, Mary A.
; APPLICANT: Pan, James
; APPLICANT: Paoni, Nicholas F.
; APPLICANT: Roy, Margaret Ann
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Watanabe, Colin K.
; APPLICANT: Williams, P. Mickey
; APPLICANT: Wood, William I.
; APPLICANT: Zhang, Zemin

; TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
; TITLE OF INVENTION: Acids Encoding the Same

; FILE REFERENCE: P2730PIC57

; CURRENT APPLICATION NUMBER: US/09/989,732

; CURRENT FILING DATE: 2001-11-19

; PRIOR APPLICATION NUMBER: 60/049787

; PRIOR FILING DATE: 1997-06-16

; PRIOR APPLICATION NUMBER: 60/062250

; PRIOR FILING DATE: 1997-10-17

; PRIOR APPLICATION NUMBER: 60/065186

; PRIOR FILING DATE: 1997-11-12

; PRIOR APPLICATION NUMBER: 60/065311

; PRIOR FILING DATE: 1997-11-13

; PRIOR APPLICATION NUMBER: 60/066770

; PRIOR FILING DATE: 1997-11-24

; PRIOR APPLICATION NUMBER: 60/075945

; PRIOR FILING DATE: 1998-02-25

; PRIOR APPLICATION NUMBER: 60/078910

; PRIOR FILING DATE: 1998-03-20

; PRIOR APPLICATION NUMBER: 60/083322

; PRIOR FILING DATE: 1998-04-28

; PRIOR APPLICATION NUMBER: 60/084600

; PRIOR FILING DATE: 1998-05-07

; PRIOR APPLICATION NUMBER: 60/087106

; PRIOR FILING DATE: 1998-05-28

; PRIOR APPLICATION NUMBER: 60/087607

; PRIOR FILING DATE: 1998-06-02

; PRIOR APPLICATION NUMBER: 60/087609

; PRIOR FILING DATE: 1998-06-02

; PRIOR APPLICATION NUMBER: 60/087759

; PRIOR FILING DATE: 1998-06-02

; PRIOR APPLICATION NUMBER: 60/087827

; PRIOR FILING DATE: 1998-06-03

09-833799-13b.rnpb

Sun Feb 16 09:13:26 2003

;
; PRIOR APPLICATION NUMBER: 60/088021
; PRIOR FILING DATE: 1998-06-04
; PRIOR APPLICATION NUMBER: 60/088025
; PRIOR FILING DATE: 1998-06-04
; PRIOR APPLICATION NUMBER: 60/088026
; PRIOR FILING DATE: 1998-06-04
; PRIOR APPLICATION NUMBER: 60/088028
; PRIOR FILING DATE: 1998-06-04
; PRIOR APPLICATION NUMBER: 60/088029
; PRIOR FILING DATE: 1998-06-04
; PRIOR APPLICATION NUMBER: 60/088030
; PRIOR FILING DATE: 1998-06-04
; PRIOR APPLICATION NUMBER: 60/088033
; PRIOR FILING DATE: 1998-06-04
; PRIOR APPLICATION NUMBER: 60/088326
; PRIOR FILING DATE: 1998-06-04
; PRIOR APPLICATION NUMBER: 60/088167
; PRIOR FILING DATE: 1998-06-05
; PRIOR APPLICATION NUMBER: 60/088202
; PRIOR FILING DATE: 1998-06-05
; PRIOR APPLICATION NUMBER: 60/088212
; PRIOR FILING DATE: 1998-06-05
; PRIOR APPLICATION NUMBER: 60/088217
; PRIOR FILING DATE: 1998-06-05
; PRIOR APPLICATION NUMBER: 60/088655
; PRIOR FILING DATE: 1998-06-09
; PRIOR APPLICATION NUMBER: 60/088734
; PRIOR FILING DATE: 1998-06-10
; PRIOR APPLICATION NUMBER: 60/088738
; PRIOR FILING DATE: 1998-06-10
; PRIOR APPLICATION NUMBER: 60/088742
; PRIOR FILING DATE: 1998-06-10
; PRIOR APPLICATION NUMBER: 60/088810
; PRIOR FILING DATE: 1998-06-10
; PRIOR APPLICATION NUMBER: 60/088824
; PRIOR FILING DATE: 1998-06-10
; PRIOR APPLICATION NUMBER: 60/088826
; PRIOR FILING DATE: 1998-06-10
; PRIOR APPLICATION NUMBER: 60/088858
; PRIOR FILING DATE: 1998-06-11
; PRIOR APPLICATION NUMBER: 60/088861
; PRIOR FILING DATE: 1998-06-11
; PRIOR APPLICATION NUMBER: 60/088876
; PRIOR FILING DATE: 1998-06-11
; PRIOR APPLICATION NUMBER: 60/089105
; PRIOR FILING DATE: 1998-06-12
; PRIOR APPLICATION NUMBER: 60/089440
; PRIOR FILING DATE: 1998-06-16
; PRIOR APPLICATION NUMBER: 60/089512
; PRIOR FILING DATE: 1998-06-16
; PRIOR APPLICATION NUMBER: 60/089514
; PRIOR FILING DATE: 1998-06-16
; PRIOR APPLICATION NUMBER: 60/089532
; PRIOR FILING DATE: 1998-06-17
; PRIOR APPLICATION NUMBER: 60/089538
; PRIOR FILING DATE: 1998-06-17
; PRIOR APPLICATION NUMBER: 60/089598
; PRIOR FILING DATE: 1998-06-17
; PRIOR APPLICATION NUMBER: 60/089599
; PRIOR FILING DATE: 1998-06-17
; PRIOR APPLICATION NUMBER: 60/089600
; PRIOR FILING DATE: 1998-06-17
; PRIOR APPLICATION NUMBER: 60/089653
; PRIOR FILING DATE: 1998-06-17
; PRIOR APPLICATION NUMBER: 60/089801
; PRIOR FILING DATE: 1998-06-18
; PRIOR APPLICATION NUMBER: 60/089907
; PRIOR FILING DATE: 1998-06-18
; PRIOR APPLICATION NUMBER: 60/089908
; PRIOR FILING DATE: 1998-06-18
; PRIOR APPLICATION NUMBER: 60/089947
; PRIOR FILING DATE: 1998-06-19
; PRIOR APPLICATION NUMBER: 60/089948

;
; PRIOR FILING DATE: 1998-06-19
; PRIOR APPLICATION NUMBER: 60/089952
; PRIOR FILING DATE: 1998-06-19
; PRIOR APPLICATION NUMBER: 60/090246
; PRIOR FILING DATE: 1998-06-22
; PRIOR APPLICATION NUMBER: 60/090252
; PRIOR FILING DATE: 1998-06-22
; PRIOR APPLICATION NUMBER: 60/090254
; PRIOR FILING DATE: 1998-06-22
; PRIOR APPLICATION NUMBER: 60/090349
; PRIOR FILING DATE: 1998-06-23
; PRIOR APPLICATION NUMBER: 60/090355
; PRIOR FILING DATE: 1998-06-23
; PRIOR APPLICATION NUMBER: 60/090429
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090431
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090435
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090444
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090445
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090472
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090535
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090540
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090542
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090557
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090676
; PRIOR FILING DATE: 1998-06-25
; PRIOR APPLICATION NUMBER: 60/090678
; PRIOR FILING DATE: 1998-06-25
; PRIOR APPLICATION NUMBER: 60/090690
; PRIOR FILING DATE: 1998-06-25
; PRIOR APPLICATION NUMBER: 60/090694
; PRIOR FILING DATE: 1998-06-25
; PRIOR APPLICATION NUMBER: 60/090695
; PRIOR FILING DATE: 1998-06-25
; PRIOR APPLICATION NUMBER: 60/090696
; PRIOR FILING DATE: 1998-06-25
; PRIOR APPLICATION NUMBER: 60/090862
; PRIOR FILING DATE: 1998-06-26
; PRIOR APPLICATION NUMBER: 60/090863
; PRIOR FILING DATE: 1998-06-26
; PRIOR APPLICATION NUMBER: 60/091360
; PRIOR FILING DATE: 1998-07-01
; PRIOR APPLICATION NUMBER: 60/091478
; PRIOR FILING DATE: 1998-07-02
; PRIOR APPLICATION NUMBER: 60/091544
; PRIOR FILING DATE: 1998-07-01
; PRIOR APPLICATION NUMBER: 60/091519
; PRIOR FILING DATE: 1998-07-02
; PRIOR APPLICATION NUMBER: 60/091626
; PRIOR FILING DATE: 1998-07-02
; PRIOR APPLICATION NUMBER: 60/091633
; PRIOR FILING DATE: 1998-07-02
; PRIOR APPLICATION NUMBER: 60/091978
; PRIOR FILING DATE: 1998-07-07
; PRIOR APPLICATION NUMBER: 60/091982
; PRIOR FILING DATE: 1998-07-07
; PRIOR APPLICATION NUMBER: 60/092182
; PRIOR FILING DATE: 1998-07-09

Alignment Scores:
Pred. No.: 1.9e-05
Score: 112.00
Percent Similarity: 53.57%
Best Local Similarity: 42.86%

Length: 762
Matches: 24
Conservative: 6
Mismatch: 24

```
Query Match: 34.25% Indels: 2
DB: 10 Gaps: 2
09-833799-13B (1-57) x US-09-989-732-344 (1-762)
QY 1 AlaGlnGluProValLysGlyProValSerThrLysProGlySerCysProIleLeu 20
Db 62 GCTGTGAAGAGGTTAAAGAGGGTATAGAG---AAAGCAGGGGTTTGGCCAGCTGACAAC 118
QY 21 IleArgCysAlaMetLeuAsnProProAsnArgCysLeuLysAspThrAspCysProGly 40
Db 119 GTACGCTGCTTCAAGTCGATCCTCCC---CAGTGTACACAGACCAGGACTGTCTGGGG 175
QY 41 IleLysLysCysCysGluGlySerCysGlyMetAlaCysPheValPro 56
Db 176 GAAAGGAAGTGTGTGTACTGTCACCTGTGGCTTCAAGTGTGTGATTCCT 223
RESULT 29
US-09-991-073-344
; Sequence 344, Application US/09991073
; Patent No. US20020127576A1
; GENERAL INFORMATION:
; APPLICANT: Ashkenazi, Avi J.
; APPLICANT: Baker, Kevin P.
; APPLICANT: Botstein, David
; APPLICANT: Desnoyers, Luc
; APPLICANT: Eaton, Dan L.
; APPLICANT: Ferrara, Napoleone
; APPLICANT: Fong, Sherman
; APPLICANT: Gerber, Hanspeter
; APPLICANT: Gerritsen, Mary E.
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Grimaldi, J. Christopher
; APPLICANT: Gurney, Austin L.
; APPLICANT: Kljavin, Ivar J.
; APPLICANT: Napier, Mary A.
; APPLICANT: Pan, James
; APPLICANT: Paoni, Nicholas F.
; APPLICANT: Roy, Margaret Ann
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Watanabe, Colin K.
; APPLICANT: Williams, P. Mickey
; APPLICANT: Wood, William I.
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
; FILE REFERENCE: P2730P1C15
; CURRENT APPLICATION NUMBER: US/09/991,073
; CURRENT FILING DATE: 2001-11-14
; PRIOR APPLICATION NUMBER: 60/049787
; PRIOR FILING DATE: 1997-06-16
; PRIOR APPLICATION NUMBER: 60/062250
; PRIOR FILING DATE: 1997-10-17
; PRIOR APPLICATION NUMBER: 60/065186
; PRIOR FILING DATE: 1997-11-12
; PRIOR APPLICATION NUMBER: 60/065311
; PRIOR FILING DATE: 1997-11-13
; PRIOR APPLICATION NUMBER: 60/066770
; PRIOR FILING DATE: 1997-11-24
; PRIOR APPLICATION NUMBER: 60/075945
; PRIOR FILING DATE: 1998-02-25
; PRIOR APPLICATION NUMBER: 60/078910
; PRIOR FILING DATE: 1998-03-20
; PRIOR APPLICATION NUMBER: 60/083322
; PRIOR FILING DATE: 1998-04-28
; PRIOR APPLICATION NUMBER: 60/084600
; PRIOR FILING DATE: 1998-05-07
; PRIOR APPLICATION NUMBER: 60/087106
; PRIOR FILING DATE: 1998-05-28
; PRIOR APPLICATION NUMBER: 60/087607
; PRIOR FILING DATE: 1998-06-02
; PRIOR APPLICATION NUMBER: 60/087609
; PRIOR FILING DATE: 1998-06-02
; PRIOR APPLICATION NUMBER: 60/087759
; PRIOR FILING DATE: 1998-06-02
; PRIOR APPLICATION NUMBER: 60/087827
; PRIOR FILING DATE: 1998-06-03
; PRIOR APPLICATION NUMBER: 60/088021
; PRIOR FILING DATE: 1998-06-04
; PRIOR APPLICATION NUMBER: 60/088025
; PRIOR FILING DATE: 1998-06-04
; PRIOR APPLICATION NUMBER: 60/088026
; PRIOR FILING DATE: 1998-06-04
; PRIOR APPLICATION NUMBER: 60/088028
; PRIOR FILING DATE: 1998-06-04
; PRIOR APPLICATION NUMBER: 60/088029
; PRIOR FILING DATE: 1998-06-04
; PRIOR APPLICATION NUMBER: 60/088030
; PRIOR FILING DATE: 1998-06-04
; PRIOR APPLICATION NUMBER: 60/088033
; PRIOR FILING DATE: 1998-06-04
; PRIOR APPLICATION NUMBER: 60/088326
; PRIOR FILING DATE: 1998-06-04
; PRIOR APPLICATION NUMBER: 60/088167
; PRIOR FILING DATE: 1998-06-05
; PRIOR APPLICATION NUMBER: 60/088202
; PRIOR FILING DATE: 1998-06-05
; PRIOR APPLICATION NUMBER: 60/088212
; PRIOR FILING DATE: 1998-06-05
; PRIOR APPLICATION NUMBER: 60/088217
; PRIOR FILING DATE: 1998-06-05
; PRIOR APPLICATION NUMBER: 60/088655
; PRIOR FILING DATE: 1998-06-09
; PRIOR APPLICATION NUMBER: 60/088734
; PRIOR FILING DATE: 1998-06-10
; PRIOR APPLICATION NUMBER: 60/088738
; PRIOR FILING DATE: 1998-06-10
; PRIOR APPLICATION NUMBER: 60/088742
; PRIOR FILING DATE: 1998-06-10
; PRIOR APPLICATION NUMBER: 60/088810
; PRIOR FILING DATE: 1998-06-10
; PRIOR APPLICATION NUMBER: 60/088824
; PRIOR FILING DATE: 1998-06-10
; PRIOR APPLICATION NUMBER: 60/088826
; PRIOR FILING DATE: 1998-06-10
; PRIOR APPLICATION NUMBER: 60/088858
; PRIOR FILING DATE: 1998-06-11
; PRIOR APPLICATION NUMBER: 60/088861
; PRIOR FILING DATE: 1998-06-11
; PRIOR APPLICATION NUMBER: 60/088876
; PRIOR FILING DATE: 1998-06-11
; PRIOR APPLICATION NUMBER: 60/089105
; PRIOR FILING DATE: 1998-06-12
; PRIOR APPLICATION NUMBER: 60/089440
; PRIOR FILING DATE: 1998-06-16
; PRIOR APPLICATION NUMBER: 60/089512
; PRIOR FILING DATE: 1998-06-16
; PRIOR APPLICATION NUMBER: 60/089514
; PRIOR FILING DATE: 1998-06-16
; PRIOR APPLICATION NUMBER: 60/089532
; PRIOR FILING DATE: 1998-06-17
; PRIOR APPLICATION NUMBER: 60/089538
; PRIOR FILING DATE: 1998-06-17
; PRIOR APPLICATION NUMBER: 60/089598
; PRIOR FILING DATE: 1998-06-17
; PRIOR APPLICATION NUMBER: 60/089599
; PRIOR FILING DATE: 1998-06-17
; PRIOR APPLICATION NUMBER: 60/089600
; PRIOR FILING DATE: 1998-06-17
; PRIOR APPLICATION NUMBER: 60/089653
; PRIOR FILING DATE: 1998-06-17
; PRIOR APPLICATION NUMBER: 60/089801
; PRIOR FILING DATE: 1998-06-18
; PRIOR APPLICATION NUMBER: 60/089907
```

Alignment Scores:					
Pred. No.:	1.9e-05	Length:	762		
Score:	112.00	Matches:	24		
Percent Similarity:	53.57%	Conservative:	6		
Best Local Similarity:	42.86%	Mismatches:	24		
Query Match:	34.25%	Indels:	2		
DB:	10	Gaps:	2		
09-833799-13B (1-57) x US-09-991-073-344 (1-762)					
QY 1 AlaGlnGluProValLysGlyProValSerThrLysProGlySerCysProIleLeu 20			:::		
Db 62 GCTGTGAAGGAGTTAAACAGGGTATAGAG---AAACAGGGGTTTGCCACGCTGACAAC 118					
QY 21 IleArgCysAlaMetLeuAsnProAsnArgCysLeuLysAspThrAspCysProGly 40	:::	:::	:::	:::	:::
Db 119 GTACGCTGCTTCAAGTCGCATCCTCCC---CAGTGTACACACAGACCAGGACTGTCTGGG 175					
QY 41 IleLysLysCysCysGluGlySerCysGlyMetAlaCysPheValPro 56	:::	:::	:::	:::	:::
176 CAACCAACTCTTGTATTACCTGCACCTGTGGCTTCAAGTGTGTGATTCCT 223					

```

RESULT 30
US-09-990-442-344
; Sequence 344, Application US/09990442
; Patent No. US2002032252A1
; GENERAL INFORMATION:
; APPLICANT: Ashkenazi, Avi J.
; APPLICANT: Baker, Kevin P.
; APPLICANT: Botstein, David
; APPLICANT: Desnoyers, Luc
; APPLICANT: Eaton, Dan L.
; APPLICANT: Ferrara, Napoleone
; APPLICANT: Fong, Sherman
; APPLICANT: Gerber, Hanspeter
; APPLICANT: Gerritsen, Mary E.
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Grimaldi, J. Christopher
; APPLICANT: Gurney, Austin L.
; APPLICANT: Kljavin, Ivar J.
; APPLICANT: Napier, Mary A.
; APPLICANT: Pan, James
; APPLICANT: Paoni, Nicholas F.
; APPLICANT: Roy, Margaret Ann
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Watanabe, Colin K.
; APPLICANT: Williams, P. Mickey
; APPLICANT: Wood, William I.
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: Secreted and Tra
; TITLE OF INVENTION: Acids Encoding
; FILE REFERENCE: P2730PIC8
; CURRENT APPLICATION NUMBER: US/09/99
; CURRENT FILING DATE: 2001-11-14
; PRIOR APPLICATION NUMBER: 60/049787
; PRIOR FILING DATE: 1997-06-16
; PRIOR APPLICATION NUMBER: 60/062250
; PRIOR FILING DATE: 1997-10-17
; PRIOR APPLICATION NUMBER: 60/065186
; PRIOR FILING DATE: 1997-11-12
; PRIOR APPLICATION NUMBER: 60/065311
; PRIOR FILING DATE: 1997-11-13
; PRIOR APPLICATION NUMBER: 60/066770
; PRIOR FILING DATE: 1997-11-24
; PRIOR APPLICATION NUMBER: 60/075945
; PRIOR FILING DATE: 1998-02-25
; PRIOR APPLICATION NUMBER: 60/078910
; PRIOR FILING DATE: 1998-03-20
; PRIOR APPLICATION NUMBER: 60/083322
; PRIOR FILING DATE: 1998-04-28

```

;	PRIOR FILING DATE: 1998-06-17	
;	PRIOR APPLICATION NUMBER: 60/089653	
;	PRIOR FILING DATE: 1998-06-17	
;	PRIOR APPLICATION NUMBER: 60/089801	
;	PRIOR FILING DATE: 1998-06-18	
;	PRIOR APPLICATION NUMBER: 60/089907	
;	PRIOR FILING DATE: 1998-06-18	
;	PRIOR APPLICATION NUMBER: 60/089908	
;	PRIOR FILING DATE: 1998-06-18	
;	PRIOR APPLICATION NUMBER: 60/089947	
;	PRIOR FILING DATE: 1998-06-19	
;	PRIOR APPLICATION NUMBER: 60/089948	
;	PRIOR FILING DATE: 1998-06-19	
;	PRIOR APPLICATION NUMBER: 60/089952	
;	PRIOR FILING DATE: 1998-06-19	
;	PRIOR APPLICATION NUMBER: 60/090246	
;	PRIOR FILING DATE: 1998-06-22	
;	PRIOR APPLICATION NUMBER: 60/090252	
;	PRIOR FILING DATE: 1998-06-22	
;	PRIOR APPLICATION NUMBER: 60/090254	
;	PRIOR FILING DATE: 1998-06-22	
;	PRIOR APPLICATION NUMBER: 60/090349	
;	PRIOR FILING DATE: 1998-06-23	
;	PRIOR APPLICATION NUMBER: 60/090355	
;	PRIOR FILING DATE: 1998-06-23	
;	PRIOR APPLICATION NUMBER: 60/090429	
;	PRIOR FILING DATE: 1998-06-24	
;	PRIOR APPLICATION NUMBER: 60/090431	
;	PRIOR FILING DATE: 1998-06-24	
;	PRIOR APPLICATION NUMBER: 60/090435	
;	PRIOR FILING DATE: 1998-06-24	
;	PRIOR APPLICATION NUMBER: 60/090444	
;	PRIOR FILING DATE: 1998-06-24	
;	PRIOR APPLICATION NUMBER: 60/090445	
;	PRIOR FILING DATE: 1998-06-24	
;	PRIOR APPLICATION NUMBER: 60/090472	
;	PRIOR FILING DATE: 1998-06-24	
;	PRIOR APPLICATION NUMBER: 60/090535	
;	PRIOR FILING DATE: 1998-06-24	
;	PRIOR APPLICATION NUMBER: 60/090540	
;	PRIOR FILING DATE: 1998-06-24	
;	PRIOR APPLICATION NUMBER: 60/090542	
;	PRIOR FILING DATE: 1998-06-24	
;	PRIOR APPLICATION NUMBER: 60/090557	
;	PRIOR FILING DATE: 1998-06-24	
;	PRIOR APPLICATION NUMBER: 60/090676	
;	PRIOR FILING DATE: 1998-06-25	
;	PRIOR APPLICATION NUMBER: 60/090678	
;	PRIOR FILING DATE: 1998-06-25	
;	PRIOR APPLICATION NUMBER: 60/090690	
;	PRIOR FILING DATE: 1998-06-25	
;	PRIOR APPLICATION NUMBER: 60/090694	
;	PRIOR FILING DATE: 1998-06-25	
;	PRIOR APPLICATION NUMBER: 60/090695	
;	PRIOR FILING DATE: 1998-06-25	
;	PRIOR APPLICATION NUMBER: 60/090696	
;	PRIOR FILING DATE: 1998-06-25	
;	PRIOR APPLICATION NUMBER: 60/090862	
;	PRIOR FILING DATE: 1998-06-26	
;	PRIOR APPLICATION NUMBER: 60/090863	
;	PRIOR FILING DATE: 1998-06-26	
;	PRIOR APPLICATION NUMBER: 60/091360	
;	PRIOR FILING DATE: 1998-07-01	
;	PRIOR APPLICATION NUMBER: 60/091478	
;	PRIOR FILING DATE: 1998-07-02	
;	PRIOR APPLICATION NUMBER: 60/091544	
;	PRIOR FILING DATE: 1998-07-01	
;	PRIOR APPLICATION NUMBER: 60/091519	
;	PRIOR FILING DATE: 1998-07-02	
;	PRIOR APPLICATION NUMBER: 60/091626	
;	PRIOR FILING DATE: 1998-07-02	
;	PRIOR APPLICATION NUMBER: 60/091633	
;	PRIOR FILING DATE: 1998-07-02	

Sun Feb 16 09:13:26 2003

09-833799-13b.rnpb

; PRIOR APPLICATION NUMBER: 60/091978
; PRIOR FILING DATE: 1998-07-07
; PRIOR APPLICATION NUMBER: 60/091982
; PRIOR FILING DATE: 1998-07-07
; PRIOR APPLICATION NUMBER: 60/092182
; PRIOR FILING DATE: 1998-07-09

Alignment Scores:
Pred. No.: 1.9e-05 Length: 762
Score: 112.00 Matches: 24
Percent Similarity: 53.57% Conservative: 6
Best Local Similarity: 42.86% Mismatches: 24
Query Match: 34.25% Indels: 2
DB: 10 Gaps: 2

09-833799-13B (1-57) x US-09-990-442-344 (1-762)

QY 1 AlaGlnGluProValLysGlyProValSerThrLysProGlySerCysProIleLeu 20
Db 62 GCTGTGAAGGAGTTAAAGAGGGTATAGAG---AAAGCAGGGGTTTGCCCGAGCTGACAAC 118
QY 21 IleArgCysAlaMetLeuAsnProProAsnArgCysLeuLysAspThrAspCysProGly 40
Db 119 GTACGCTGCTTCAAGTCCGATCTCTCC---CAGTGTCAACAGACAGGAGCTGTCTGGG 175
QY 41 IleLysLysCysCysGluGlySerCysGlyMetAlaCysPheValPro 56
Db 176 GAAAGGAAGTGTGTTACCTGCACTGTGGCTTCAAGTGTGTGATTCCT 223

RESULT 31

US-09-991-163-344
Sequence 344, Application US/09991163

; Patent No. US20020132253A1

; GENERAL INFORMATION:

; APPLICANT: Ashkenazi, Avi J.
; APPLICANT: Baker, Kevin P.
; APPLICANT: Botstein, David
; APPLICANT: Desnoyers, Luc
; APPLICANT: Eaton, Dan L.
; APPLICANT: Ferrara, Napoleone
; APPLICANT: Fong, Sherman
; APPLICANT: Gerber, Hanspeter
; APPLICANT: Gerritsen, Mary E.
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Grimaldi, J. Christopher
; APPLICANT: Gurney, Austin L.
; APPLICANT: Kljavin, Ivar J.
; APPLICANT: Napier, Mary A.
; APPLICANT: Pan, James
; APPLICANT: Paoni, Nicholas F.
; APPLICANT: Roy, Margaret Ann
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Watanabe, Colin K.
; APPLICANT: Williams, P. Mickey
; APPLICANT: Wood, William I.
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
; TITLE OF INVENTION: Acids Encoding the Same
; FILE REFERENCE: P2730P1C17
; CURRENT APPLICATION NUMBER: US/09/991,163
; CURRENT FILING DATE: 2001-11-14
; PRIOR APPLICATION NUMBER: 60/049787
; PRIOR FILING DATE: 1997-06-16
; PRIOR APPLICATION NUMBER: 60/062250
; PRIOR FILING DATE: 1997-10-17
; PRIOR APPLICATION NUMBER: 60/065186
; PRIOR FILING DATE: 1997-11-12
; PRIOR APPLICATION NUMBER: 60/065311
; PRIOR FILING DATE: 1997-11-13
; PRIOR APPLICATION NUMBER: 60/066770
; PRIOR FILING DATE: 1997-11-24

; PRIOR APPLICATION NUMBER: 60/075945
; PRIOR FILING DATE: 1998-02-25
; PRIOR APPLICATION NUMBER: 60/078910
; PRIOR FILING DATE: 1998-03-20
; PRIOR APPLICATION NUMBER: 60/083322
; PRIOR FILING DATE: 1998-04-28
; PRIOR APPLICATION NUMBER: 60/084600
; PRIOR FILING DATE: 1998-05-07
; PRIOR APPLICATION NUMBER: 60/087106
; PRIOR FILING DATE: 1998-05-28
; PRIOR APPLICATION NUMBER: 60/087607
; PRIOR FILING DATE: 1998-06-02
; PRIOR APPLICATION NUMBER: 60/087609
; PRIOR FILING DATE: 1998-06-02
; PRIOR APPLICATION NUMBER: 60/087759
; PRIOR FILING DATE: 1998-06-02
; PRIOR APPLICATION NUMBER: 60/087827
; PRIOR FILING DATE: 1998-06-03
; PRIOR APPLICATION NUMBER: 60/088021
; PRIOR FILING DATE: 1998-06-04
; PRIOR APPLICATION NUMBER: 60/088025
; PRIOR FILING DATE: 1998-06-04
; PRIOR APPLICATION NUMBER: 60/088026
; PRIOR FILING DATE: 1998-06-04
; PRIOR APPLICATION NUMBER: 60/088028
; PRIOR FILING DATE: 1998-06-04
; PRIOR APPLICATION NUMBER: 60/088029
; PRIOR FILING DATE: 1998-06-04
; PRIOR APPLICATION NUMBER: 60/088030
; PRIOR FILING DATE: 1998-06-04
; PRIOR APPLICATION NUMBER: 60/088033
; PRIOR FILING DATE: 1998-06-04
; PRIOR APPLICATION NUMBER: 60/088326
; PRIOR FILING DATE: 1998-06-04
; PRIOR APPLICATION NUMBER: 60/088167
; PRIOR FILING DATE: 1998-06-05
; PRIOR APPLICATION NUMBER: 60/088202
; PRIOR FILING DATE: 1998-06-05
; PRIOR APPLICATION NUMBER: 60/088212
; PRIOR FILING DATE: 1998-06-05
; PRIOR APPLICATION NUMBER: 60/088217
; PRIOR FILING DATE: 1998-06-05
; PRIOR APPLICATION NUMBER: 60/088655
; PRIOR FILING DATE: 1998-06-09
; PRIOR APPLICATION NUMBER: 60/088734
; PRIOR FILING DATE: 1998-06-10
; PRIOR APPLICATION NUMBER: 60/088738
; PRIOR FILING DATE: 1998-06-10
; PRIOR APPLICATION NUMBER: 60/088742
; PRIOR FILING DATE: 1998-06-10
; PRIOR APPLICATION NUMBER: 60/088810
; PRIOR FILING DATE: 1998-06-10
; PRIOR APPLICATION NUMBER: 60/088824
; PRIOR FILING DATE: 1998-06-10
; PRIOR APPLICATION NUMBER: 60/088826
; PRIOR FILING DATE: 1998-06-10
; PRIOR APPLICATION NUMBER: 60/088858
; PRIOR FILING DATE: 1998-06-11
; PRIOR APPLICATION NUMBER: 60/088861
; PRIOR FILING DATE: 1998-06-11
; PRIOR APPLICATION NUMBER: 60/088876
; PRIOR FILING DATE: 1998-06-11
; PRIOR APPLICATION NUMBER: 60/089105
; PRIOR FILING DATE: 1998-06-12
; PRIOR APPLICATION NUMBER: 60/089440
; PRIOR FILING DATE: 1998-06-16
; PRIOR APPLICATION NUMBER: 60/089512
; PRIOR FILING DATE: 1998-06-16
; PRIOR APPLICATION NUMBER: 60/089514
; PRIOR FILING DATE: 1998-06-16
; PRIOR APPLICATION NUMBER: 60/089532
; PRIOR FILING DATE: 1998-06-17
; PRIOR APPLICATION NUMBER: 60/089538

;
; PRIOR FILING DATE: 1998-06-17
; PRIOR APPLICATION NUMBER: 60/089598
; PRIOR FILING DATE: 1998-06-17
; PRIOR APPLICATION NUMBER: 60/089599
; PRIOR FILING DATE: 1998-06-17
; PRIOR APPLICATION NUMBER: 60/089600
; PRIOR FILING DATE: 1998-06-17
; PRIOR APPLICATION NUMBER: 60/089653
; PRIOR FILING DATE: 1998-06-17
; PRIOR APPLICATION NUMBER: 60/089801
; PRIOR FILING DATE: 1998-06-18
; PRIOR APPLICATION NUMBER: 60/089907
; PRIOR FILING DATE: 1998-06-18
; PRIOR APPLICATION NUMBER: 60/089908
; PRIOR FILING DATE: 1998-06-18
; PRIOR APPLICATION NUMBER: 60/089947
; PRIOR FILING DATE: 1998-06-19
; PRIOR APPLICATION NUMBER: 60/089948
; PRIOR FILING DATE: 1998-06-19
; PRIOR APPLICATION NUMBER: 60/089952
; PRIOR FILING DATE: 1998-06-19
; PRIOR APPLICATION NUMBER: 60/090246
; PRIOR FILING DATE: 1998-06-22
; PRIOR APPLICATION NUMBER: 60/090252
; PRIOR FILING DATE: 1998-06-22
; PRIOR APPLICATION NUMBER: 60/090254
; PRIOR FILING DATE: 1998-06-22
; PRIOR APPLICATION NUMBER: 60/090349
; PRIOR FILING DATE: 1998-06-23
; PRIOR APPLICATION NUMBER: 60/090355
; PRIOR FILING DATE: 1998-06-23
; PRIOR APPLICATION NUMBER: 60/090429
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090431
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090435
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090444
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090445
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090472
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090535
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090540
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090542
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090557
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090676
; PRIOR FILING DATE: 1998-06-25
; PRIOR APPLICATION NUMBER: 60/090678
; PRIOR FILING DATE: 1998-06-25
; PRIOR APPLICATION NUMBER: 60/090690
; PRIOR FILING DATE: 1998-06-25
; PRIOR APPLICATION NUMBER: 60/090694
; PRIOR FILING DATE: 1998-06-25
; PRIOR APPLICATION NUMBER: 60/090695
; PRIOR FILING DATE: 1998-06-25
; PRIOR APPLICATION NUMBER: 60/090696
; PRIOR FILING DATE: 1998-06-25
; PRIOR APPLICATION NUMBER: 60/090862
; PRIOR FILING DATE: 1998-06-26
; PRIOR APPLICATION NUMBER: 60/090863
; PRIOR FILING DATE: 1998-06-26
; PRIOR APPLICATION NUMBER: 60/091360
; PRIOR FILING DATE: 1998-07-01
; PRIOR APPLICATION NUMBER: 60/091478
; PRIOR FILING DATE: 1998-07-02
; PRIOR APPLICATION NUMBER: 60/091544
; PRIOR FILING DATE: 1998-07-01

;
; PRIOR APPLICATION NUMBER: 60/091519
; PRIOR FILING DATE: 1998-07-02
; PRIOR APPLICATION NUMBER: 60/091626
; PRIOR FILING DATE: 1998-07-02
; PRIOR APPLICATION NUMBER: 60/091633
; PRIOR FILING DATE: 1998-07-02
; PRIOR APPLICATION NUMBER: 60/091978
; PRIOR FILING DATE: 1998-07-07
; PRIOR APPLICATION NUMBER: 60/091982
; PRIOR FILING DATE: 1998-07-07
; PRIOR APPLICATION NUMBER: 60/092182
; PRIOR FILING DATE: 1998-07-09

Alignment Scores:

Pred. No.:	1.9e-05	Length:	762
Score:	112.00	Matches:	24
Percent Similarity:	53.57%	Conservative:	6
Best Local Similarity:	42.86%	Mismatches:	24
Query Match:	34.25%	Indels:	2
DB:	10	Gaps:	2

09-833799-13B (1-57) x US-09-991-163-344 (1-762)

QY	1	AlaGlnGluProValLysGlyProValSerThrLysProGlySerCysProIleLeu	20
Db	62	GCTGTGAAGGAGTTAAACAGGGTATAGAG---AAAGCAGGGGTTTGGCCAGCTGACAAC	118
QY	21	IleArgCysAlaMetLeuAsnProAsnArgCysLeuLysAspThrAspCysProGly	40
Db	119	GTACGCTGCTTCAAGTCCGATCCTCCC---CAGTGTACACAGACCAGGACTGTCTGGGG	175
QY	41	IleLysLysCysCysGluGlySerCysGlyMetAlaCysPheValPro	56
Db	176	GAAAGGAAGTGTGTGTTACTGCACTGTGGCTTCAAGTGTGTGATTCCT	223

RESULT 32

US-09-993-604-344
; Sequence 344, Application US/09993604
; Patent No. US20020137075A1
; GENERAL INFORMATION:
; APPLICANT: Ashkenazi, Avi J.
; APPLICANT: Baker, Kevin P.
; APPLICANT: Botstein, David
; APPLICANT: Desnoyers, Luc
; APPLICANT: Eaton, Dan L.
; APPLICANT: Ferrara, Napoleone
; APPLICANT: Fong, Sherman
; APPLICANT: Gerber, Hanspeter
; APPLICANT: Gerritsen, Mary E.
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Grimaldi, J. Christopher
; APPLICANT: Gurney, Austin L.
; APPLICANT: Kljavin, Ivar J.
; APPLICANT: Napier, Mary A.
; APPLICANT: Pan, James
; APPLICANT: Paoni, Nicholas F.
; APPLICANT: Roy, Margaret Ann
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Watanabe, Colin K.
; APPLICANT: Williams, P. Mickey
; APPLICANT: Wood, William I.
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
; TITLE OF INVENTION: Acids Encoding the Same
; FILE REFERENCE: P2730PIC25
; CURRENT APPLICATION NUMBER: US/09/993,604
; CURRENT FILING DATE: 2001-11-14
; PRIOR APPLICATION NUMBER: 60/049787
; PRIOR FILING DATE: 1997-06-16
; PRIOR APPLICATION NUMBER: 60/062250
; PRIOR FILING DATE: 1997-10-17

;	CURRENT APPLICATION NUMBER: US/09/990,450
;	CURRENT FILING DATE: 2001-11-14
;	PRIOR APPLICATION NUMBER: 60/049787
;	PRIOR FILING DATE: 1997-06-16
;	PRIOR APPLICATION NUMBER: 60/062250
;	PRIOR FILING DATE: 1997-10-17
;	PRIOR APPLICATION NUMBER: 60/065186
;	PRIOR FILING DATE: 1997-11-12
;	PRIOR APPLICATION NUMBER: 60/065311
;	PRIOR FILING DATE: 1997-11-13
;	PRIOR APPLICATION NUMBER: 60/066770
;	PRIOR FILING DATE: 1997-11-24
;	PRIOR APPLICATION NUMBER: 60/075945
;	PRIOR FILING DATE: 1998-02-25
;	PRIOR APPLICATION NUMBER: 60/078910
;	PRIOR FILING DATE: 1998-03-20
;	PRIOR APPLICATION NUMBER: 60/083322
;	PRIOR FILING DATE: 1998-04-28
;	PRIOR APPLICATION NUMBER: 60/084600
;	PRIOR FILING DATE: 1998-05-07
;	PRIOR APPLICATION NUMBER: 60/087106
;	PRIOR FILING DATE: 1998-05-28
;	PRIOR APPLICATION NUMBER: 60/087607
;	PRIOR FILING DATE: 1998-06-02
;	PRIOR APPLICATION NUMBER: 60/087609
;	PRIOR FILING DATE: 1998-06-02
;	PRIOR APPLICATION NUMBER: 60/087759
;	PRIOR FILING DATE: 1998-06-02
;	PRIOR APPLICATION NUMBER: 60/087827
;	PRIOR FILING DATE: 1998-06-03
;	PRIOR APPLICATION NUMBER: 60/088021
;	PRIOR FILING DATE: 1998-06-04
;	PRIOR APPLICATION NUMBER: 60/088025
;	PRIOR FILING DATE: 1998-06-04
;	PRIOR APPLICATION NUMBER: 60/088026
;	PRIOR FILING DATE: 1998-06-04
;	PRIOR APPLICATION NUMBER: 60/088028
;	PRIOR FILING DATE: 1998-06-04
;	PRIOR APPLICATION NUMBER: 60/088029
;	PRIOR FILING DATE: 1998-06-04
;	PRIOR APPLICATION NUMBER: 60/088030
;	PRIOR FILING DATE: 1998-06-04
;	PRIOR APPLICATION NUMBER: 60/088033
;	PRIOR FILING DATE: 1998-06-04
;	PRIOR APPLICATION NUMBER: 60/088326
;	PRIOR FILING DATE: 1998-06-04
;	PRIOR APPLICATION NUMBER: 60/088167
;	PRIOR FILING DATE: 1998-06-05
;	PRIOR APPLICATION NUMBER: 60/088202
;	PRIOR FILING DATE: 1998-06-05
;	PRIOR APPLICATION NUMBER: 60/088212
;	PRIOR FILING DATE: 1998-06-05
;	PRIOR APPLICATION NUMBER: 60/088217
;	PRIOR FILING DATE: 1998-06-05
;	PRIOR APPLICATION NUMBER: 60/088655
;	PRIOR FILING DATE: 1998-06-09
;	PRIOR APPLICATION NUMBER: 60/088734
;	PRIOR FILING DATE: 1998-06-10
;	PRIOR APPLICATION NUMBER: 60/088738
;	PRIOR FILING DATE: 1998-06-10
;	PRIOR APPLICATION NUMBER: 60/088742
;	PRIOR FILING DATE: 1998-06-10
;	PRIOR APPLICATION NUMBER: 60/088810
;	PRIOR FILING DATE: 1998-06-10
;	PRIOR APPLICATION NUMBER: 60/088824
;	PRIOR FILING DATE: 1998-06-10
;	PRIOR APPLICATION NUMBER: 60/088826
;	PRIOR FILING DATE: 1998-06-10
;	PRIOR APPLICATION NUMBER: 60/088858
;	PRIOR FILING DATE: 1998-06-11
;	PRIOR APPLICATION NUMBER: 60/088861
;	PRIOR FILING DATE: 1998-06-11
;	PRIOR APPLICATION NUMBER: 60/088876

Pred. No.:	1.9e-05	Length:	762
Score:	112.00	Matches:	24
Percent Similarity:	53.57%	Conservative:	6
Best Local Similarity:	42.86%	Mismatches:	24
Query Match:	34.25%	Indels:	2
DE:	10	Gaps:	2

Qy	1	AlaGlnGluProValLysGlyProValSerThrLysProGlySerCysProIleLeu	20
		::	
Db	62	GCTGTGGAAGGAGTTAAAGAGGGTATAGAG--AAAGCAGGGGTTGCCAGCTGCACAAC	118
Qy	21	IleArgCysAlaMetLeuAsnProProAsnArgCysLeuLysAspThrAspCysProGly	40
		::: :::	
Db	119	GTACGCTGCTCAAGTC CGATCCCTCCC---CAGTGTCACACAGACCAGGACTGTC TGGGG	175
Qy	41	IleLysLysCysCysGluGlySerCysGlyMetAlaCysPheValPro	56
		:::	
Db	176	GAAGGAAGTGTGTTACCTGCATGTGGCTTCAAGTGTGTGATTCTT	223

; FILE REFERENCE: P2730P1C22

```

; TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
; TITLE OF INVENTION: Acids Encoding the Same
; FILE REFERENCE: P2730P1C22

```

; PRIOR FILING DATE: 1998-06-11
; PRIOR APPLICATION NUMBER: 60/089105
; PRIOR FILING DATE: 1998-06-12
; PRIOR APPLICATION NUMBER: 60/089440
; PRIOR FILING DATE: 1998-06-16
; PRIOR APPLICATION NUMBER: 60/089512
; PRIOR FILING DATE: 1998-06-16
; PRIOR APPLICATION NUMBER: 60/089514
; PRIOR FILING DATE: 1998-06-16
; PRIOR APPLICATION NUMBER: 60/089532
; PRIOR FILING DATE: 1998-06-17
; PRIOR APPLICATION NUMBER: 60/089538
; PRIOR FILING DATE: 1998-06-17
; PRIOR APPLICATION NUMBER: 60/089598
; PRIOR FILING DATE: 1998-06-17
; PRIOR APPLICATION NUMBER: 60/089599
; PRIOR FILING DATE: 1998-06-17
; PRIOR APPLICATION NUMBER: 60/089600
; PRIOR FILING DATE: 1998-06-17
; PRIOR APPLICATION NUMBER: 60/089653
; PRIOR FILING DATE: 1998-06-17
; PRIOR APPLICATION NUMBER: 60/089801
; PRIOR FILING DATE: 1998-06-18
; PRIOR APPLICATION NUMBER: 60/089907
; PRIOR FILING DATE: 1998-06-18
; PRIOR APPLICATION NUMBER: 60/089908
; PRIOR FILING DATE: 1998-06-18
; PRIOR APPLICATION NUMBER: 60/089947
; PRIOR FILING DATE: 1998-06-19
; PRIOR APPLICATION NUMBER: 60/089948
; PRIOR FILING DATE: 1998-06-19
; PRIOR APPLICATION NUMBER: 60/089952
; PRIOR FILING DATE: 1998-06-19
; PRIOR APPLICATION NUMBER: 60/090246
; PRIOR FILING DATE: 1998-06-22
; PRIOR APPLICATION NUMBER: 60/090252
; PRIOR FILING DATE: 1998-06-22
; PRIOR APPLICATION NUMBER: 60/090254
; PRIOR FILING DATE: 1998-06-22
; PRIOR APPLICATION NUMBER: 60/090349
; PRIOR FILING DATE: 1998-06-23
; PRIOR APPLICATION NUMBER: 60/090355
; PRIOR FILING DATE: 1998-06-23
; PRIOR APPLICATION NUMBER: 60/090429
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090431
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090435
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090444
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090445
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090472
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090535
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090540
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090542
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090557
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090676
; PRIOR FILING DATE: 1998-06-25
; PRIOR APPLICATION NUMBER: 60/090678
; PRIOR FILING DATE: 1998-06-25
; PRIOR APPLICATION NUMBER: 60/090690
; PRIOR FILING DATE: 1998-06-25
; PRIOR APPLICATION NUMBER: 60/090694
; PRIOR FILING DATE: 1998-06-25
; PRIOR APPLICATION NUMBER: 60/090695
; PRIOR FILING DATE: 1998-06-25

; PRIOR APPLICATION NUMBER: 60/090696
; PRIOR FILING DATE: 1998-06-25
; PRIOR APPLICATION NUMBER: 60/090862
; PRIOR FILING DATE: 1998-06-26
; PRIOR APPLICATION NUMBER: 60/090863
; PRIOR FILING DATE: 1998-06-26
; PRIOR APPLICATION NUMBER: 60/091360
; PRIOR FILING DATE: 1998-07-01
; PRIOR APPLICATION NUMBER: 60/091478
; PRIOR FILING DATE: 1998-07-02
; PRIOR APPLICATION NUMBER: 60/091544
; PRIOR FILING DATE: 1998-07-01
; PRIOR APPLICATION NUMBER: 60/091519
; PRIOR FILING DATE: 1998-07-02
; PRIOR APPLICATION NUMBER: 60/091626
; PRIOR FILING DATE: 1998-07-02
; PRIOR APPLICATION NUMBER: 60/091633
; PRIOR FILING DATE: 1998-07-02
; PRIOR APPLICATION NUMBER: 60/091978
; PRIOR FILING DATE: 1998-07-07
; PRIOR APPLICATION NUMBER: 60/091982
; PRIOR FILING DATE: 1998-07-07
; PRIOR APPLICATION NUMBER: 60/092182
; PRIOR FILING DATE: 1998-07-09

Alignment Scores:
Pred. No.: 1.9e-05 Length: 762
Score: 112.00 Matches: 24
Percent Similarity: 53.57% Conservative: 6
Best Local Similarity: 42.86% Mismatches: 24
Query Match: 34.25% Indels: 2
DB: 10 Gaps: 2

09-833799-13B (1-57) x US-09-990-456-344 (1-762)

QY 1 AlaGlnGluProValLysGlyProValSerThrLysProGlySerCysProIleIleLeu 20
Db 62 GCTGTGGAAGAGGAGTTAAAGAGGGGTATAGAG---AAAGCAGGGGTTTGGCCAGCTGACAAC 118
QY 21 IleArgCysAlaMetLeuAsnProProAsnArgCysLeuLysAspThrAspCysProGly 40
Db 119 GTACGCTGCTTCAAGTCCGATCCTCCC---CAGTGTACACACAGACCAGGACTGTCTGGGG 175
QY 41 IleLysLysCysCysGluGlySerCysGlyMetAlaCysPheValPro 56
Db 176 GAAAGGAAGTGTGTACCTGCACCTGTGGCTTCAAGTGTGTGATTCT 223

RESULT 34
US-09-989-721-344
; Sequence 344, Application US/099899721
; Patent No. US20020142961A1
; GENERAL INFORMATION:
; APPLICANT: Ashkenazi, Avi J.
; APPLICANT: Baker, Kevin P.
; APPLICANT: Botstein, David
; APPLICANT: Desnoyers, Luc
; APPLICANT: Eaton, Dan L.
; APPLICANT: Ferrara, Napoleone
; APPLICANT: Fong, Sherman
; APPLICANT: Gerber, Hanspeter
; APPLICANT: Gerritsen, Mary E.
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Grimaldi, J. Christopher
; APPLICANT: Gurney, Austin L.
; APPLICANT: Kljavin, Ivar J.
; APPLICANT: Napier, Mary A.
; APPLICANT: Pan, James
; APPLICANT: Paoni, Nicholas F.
; APPLICANT: Roy, Margaret Ann
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Watanabe, Colin K.

APPLICANT: Williams, P. Mickey
APPLICANT: Wood, William I.
APPLICANT: Zhang, Zemin
TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
TITLE OF INVENTION: Acids Encoding the Same
FILE REFERENCE: P2730P1C55
CURRENT APPLICATION NUMBER: US/09/989,721
CURRENT FILING DATE: 2001-11-19
PRIOR APPLICATION NUMBER: 60/049787
PRIOR FILING DATE: 1997-06-16
PRIOR APPLICATION NUMBER: 60/062250
PRIOR FILING DATE: 1997-10-17
PRIOR APPLICATION NUMBER: 60/065186
PRIOR FILING DATE: 1997-11-12
PRIOR APPLICATION NUMBER: 60/065311
PRIOR FILING DATE: 1997-11-13
PRIOR APPLICATION NUMBER: 60/066770
PRIOR FILING DATE: 1997-11-24
PRIOR APPLICATION NUMBER: 60/075945
PRIOR FILING DATE: 1998-02-25
PRIOR APPLICATION NUMBER: 60/078910
PRIOR FILING DATE: 1998-03-20
PRIOR APPLICATION NUMBER: 60/083322
PRIOR FILING DATE: 1998-04-28
PRIOR APPLICATION NUMBER: 60/084600
PRIOR FILING DATE: 1998-05-07
PRIOR APPLICATION NUMBER: 60/087106
PRIOR FILING DATE: 1998-05-28
PRIOR APPLICATION NUMBER: 60/087607
PRIOR FILING DATE: 1998-06-02
PRIOR APPLICATION NUMBER: 60/087609
PRIOR FILING DATE: 1998-06-02
PRIOR APPLICATION NUMBER: 60/087759
PRIOR FILING DATE: 1998-06-02
PRIOR APPLICATION NUMBER: 60/087827
PRIOR FILING DATE: 1998-06-03
PRIOR APPLICATION NUMBER: 60/088021
PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088025
PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088026
PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088028
PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088029
PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088030
PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088033
PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088326
PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088167
PRIOR FILING DATE: 1998-06-05
PRIOR APPLICATION NUMBER: 60/088202
PRIOR FILING DATE: 1998-06-05
PRIOR APPLICATION NUMBER: 60/088212
PRIOR FILING DATE: 1998-06-05
PRIOR APPLICATION NUMBER: 60/088217
PRIOR FILING DATE: 1998-06-05
PRIOR APPLICATION NUMBER: 60/088655
PRIOR FILING DATE: 1998-06-09
PRIOR APPLICATION NUMBER: 60/088734
PRIOR FILING DATE: 1998-06-10
PRIOR APPLICATION NUMBER: 60/088738
PRIOR FILING DATE: 1998-06-10
PRIOR APPLICATION NUMBER: 60/088742
PRIOR FILING DATE: 1998-06-10
PRIOR APPLICATION NUMBER: 60/088810
PRIOR FILING DATE: 1998-06-10
PRIOR APPLICATION NUMBER: 60/088824
PRIOR FILING DATE: 1998-06-10
PRIOR APPLICATION NUMBER: 60/088826

PRIOR FILING DATE: 1998-06-10
PRIOR APPLICATION NUMBER: 60/088858
PRIOR FILING DATE: 1998-06-11
PRIOR APPLICATION NUMBER: 60/088861
PRIOR FILING DATE: 1998-06-11
PRIOR APPLICATION NUMBER: 60/088876
PRIOR FILING DATE: 1998-06-11
PRIOR APPLICATION NUMBER: 60/089105
PRIOR FILING DATE: 1998-06-12
PRIOR APPLICATION NUMBER: 60/089440
PRIOR FILING DATE: 1998-06-16
PRIOR APPLICATION NUMBER: 60/089512
PRIOR FILING DATE: 1998-06-16
PRIOR APPLICATION NUMBER: 60/089514
PRIOR FILING DATE: 1998-06-16
PRIOR APPLICATION NUMBER: 60/089532
PRIOR FILING DATE: 1998-06-17
PRIOR APPLICATION NUMBER: 60/089538
PRIOR FILING DATE: 1998-06-17
PRIOR APPLICATION NUMBER: 60/089598
PRIOR FILING DATE: 1998-06-17
PRIOR APPLICATION NUMBER: 60/089599
PRIOR FILING DATE: 1998-06-17
PRIOR APPLICATION NUMBER: 60/089600
PRIOR FILING DATE: 1998-06-17
PRIOR APPLICATION NUMBER: 60/089653
PRIOR FILING DATE: 1998-06-17
PRIOR APPLICATION NUMBER: 60/089801
PRIOR FILING DATE: 1998-06-18
PRIOR APPLICATION NUMBER: 60/089907
PRIOR FILING DATE: 1998-06-18
PRIOR APPLICATION NUMBER: 60/089908
PRIOR FILING DATE: 1998-06-18
PRIOR APPLICATION NUMBER: 60/089947
PRIOR FILING DATE: 1998-06-19
PRIOR APPLICATION NUMBER: 60/089948
PRIOR FILING DATE: 1998-06-19
PRIOR APPLICATION NUMBER: 60/089952
PRIOR FILING DATE: 1998-06-19
PRIOR APPLICATION NUMBER: 60/090246
PRIOR FILING DATE: 1998-06-22
PRIOR APPLICATION NUMBER: 60/090252
PRIOR FILING DATE: 1998-06-22
PRIOR APPLICATION NUMBER: 60/090254
PRIOR FILING DATE: 1998-06-22
PRIOR APPLICATION NUMBER: 60/090349
PRIOR FILING DATE: 1998-06-23
PRIOR APPLICATION NUMBER: 60/090355
PRIOR FILING DATE: 1998-06-23
PRIOR APPLICATION NUMBER: 60/090429
PRIOR FILING DATE: 1998-06-24
PRIOR APPLICATION NUMBER: 60/090431
PRIOR FILING DATE: 1998-06-24
PRIOR APPLICATION NUMBER: 60/090435
PRIOR FILING DATE: 1998-06-24
PRIOR APPLICATION NUMBER: 60/090444
PRIOR FILING DATE: 1998-06-24
PRIOR APPLICATION NUMBER: 60/090445
PRIOR FILING DATE: 1998-06-24
PRIOR APPLICATION NUMBER: 60/090472
PRIOR FILING DATE: 1998-06-24
PRIOR APPLICATION NUMBER: 60/090535
PRIOR FILING DATE: 1998-06-24
PRIOR APPLICATION NUMBER: 60/090540
PRIOR FILING DATE: 1998-06-24
PRIOR APPLICATION NUMBER: 60/090542
PRIOR FILING DATE: 1998-06-24
PRIOR APPLICATION NUMBER: 60/090557
PRIOR FILING DATE: 1998-06-24
PRIOR APPLICATION NUMBER: 60/090676
PRIOR FILING DATE: 1998-06-25
PRIOR APPLICATION NUMBER: 60/090678
PRIOR FILING DATE: 1998-06-25

; PRIOR APPLICATION NUMBER: 60/090690
; PRIOR FILING DATE: 1998-06-25
; PRIOR APPLICATION NUMBER: 60/090694
; PRIOR FILING DATE: 1998-06-25
; PRIOR APPLICATION NUMBER: 60/090695
; PRIOR FILING DATE: 1998-06-25
; PRIOR APPLICATION NUMBER: 60/090696
; PRIOR FILING DATE: 1998-06-25
; PRIOR APPLICATION NUMBER: 60/090862
; PRIOR FILING DATE: 1998-06-26
; PRIOR APPLICATION NUMBER: 60/090863
; PRIOR FILING DATE: 1998-06-26
; PRIOR APPLICATION NUMBER: 60/091360
; PRIOR FILING DATE: 1998-07-01
; PRIOR APPLICATION NUMBER: 60/091478
; PRIOR FILING DATE: 1998-07-02
; PRIOR APPLICATION NUMBER: 60/091544
; PRIOR FILING DATE: 1998-07-01
; PRIOR APPLICATION NUMBER: 60/091519
; PRIOR FILING DATE: 1998-07-02
; PRIOR APPLICATION NUMBER: 60/091626
; PRIOR FILING DATE: 1998-07-02
; PRIOR APPLICATION NUMBER: 60/091633
; PRIOR FILING DATE: 1998-07-02
; PRIOR APPLICATION NUMBER: 60/091978
; PRIOR FILING DATE: 1998-07-07
; PRIOR APPLICATION NUMBER: 60/091982
; PRIOR FILING DATE: 1998-07-07
; PRIOR APPLICATION NUMBER: 60/092182
; PRIOR FILING DATE: 1998-07-09

Alignment Scores:
Pred. No.: 1.9e-05 Length: 762
Score: 112.00 Matches: 24
Percent Similarity: 53.57% Conservative: 6
Best Local Similarity: 42.86% Mismatches: 24
Query Match: 34.25% Indels: 2
DB: 10 Gaps: 2

09-833799-13B (1-57) x US-09-989-721-344 (1-762)

Qy 1 AlaGlnGluProValLysGlyProValSerThrLysProGlySerCysProIleLeu 20
Db 62 GCTGTGAAGGAGTTAAAGAGGGTATAGAG---AAAGCAGGGGTTGCCAGCTGACAAC 118
Qy 21 IleArgCysAlaMetLeuAsnProProAsnArgCysLeuLysAspThrAspCysProGly 40
Db 119 GTACGCTGCTTCAAGTCCGATCCTCCC---CAGTGTACACAGACAGGACTGTCTGGGG 175
Qy 41 IleLysLysCysCysGluGlySerCysGlyMetAlaCysPheValPro 56
Db 176 GAAAGGAAGTGTGTTACCTGCACCTGTGGCTTCAAGTGTGTGATTCT 223

RESULT 35
US-09-964-824A-495/c
; Sequence 495, Application US/09964824A
; Patent No. US20020102531A1
; GENERAL INFORMATION:
; APPLICANT: Horrigan, Stephen
; TITLE OF INVENTION: Cancer Gene Determination and Therapeutic Screening Using Signatu
; TITLE OF INVENTION: Sets
; FILE REFERENCE: 689290-73
; CURRENT APPLICATION NUMBER: US/09/964,824A
; PRIOR FILING DATE: 2001-09-27
; PRIOR APPLICATION NUMBER: US/60/236,033
; PRIOR FILING DATE: 2000-09-28
; PRIOR APPLICATION NUMBER: US/60/236,032
; PRIOR FILING DATE: 2000-09-28
; PRIOR APPLICATION NUMBER: US/60/236,028
; PRIOR FILING DATE: 2000-09-28
; NUMBER OF SEQ ID NOS: 583
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 495

; LENGTH: 411
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)..(411)
; OTHER INFORMATION: n=a,t,g or c
US-09-964-824A-495

Alignment Scores:
Pred. No.: 3.58e-05 Length: 411
Score: 107.00 Matches: 21
Percent Similarity: 48.00% Conservative: 3
Best Local Similarity: 42.00% Mismatches: 25
Query Match: 32.72% Indels: 1
DB: 10 Gaps: 0

09-833799-13B (1-57) x US-09-964-824A-495 (1-411)

Qy 8 ProValSerThrLysProGlySerCysPro-IleIleLeuIleArgCysAlaMetLeuAs 27
Db 341 CCAACAAGGNGNAACCCCTTGGNAAGTCCNAGTANTTNTGGCCAATGTTNGATGCTNAA 282
Qy 27 nProProAsnArgCysLeuLysAspThrAspCysProGlyIleLysLysCysCysGluG1 47
Db 281 CCCCCCAATTTCTGTGAGATGGATGGCCAGTGCACAGCGTGACTTGAAGTGTTCATGGG 222
Qy 47 ySerCysGlyMetAlaCysPheValPro 56
Db 221 CATGTGTGGGAAATCCTGCGTTTCCCT 194

RESULT 36

US-09-954-456-1987/c
; Sequence 1987, Application US/09954456
; Patent No. US20020115057A1
; GENERAL INFORMATION:
; APPLICANT: Young, Paul
; TITLE OF INVENTION: Process for Identifying Anti-Cancer Therapeutic Agents Using Canc
; TITLE OF INVENTION: Sets
; FILE REFERENCE: 689290-76
; CURRENT APPLICATION NUMBER: US/09/954,456
; CURRENT FILING DATE: 2001-09-18
; PRIOR APPLICATION NUMBER: US/60/233,617
; PRIOR FILING DATE: 2000-09-18
; PRIOR APPLICATION NUMBER: US/60/234,052
; PRIOR FILING DATE: 2000-09-20
; PRIOR APPLICATION NUMBER: US/60/234,923
; PRIOR FILING DATE: 2000-09-25
; PRIOR APPLICATION NUMBER: US/60/235,134
; PRIOR FILING DATE: 2000-09-25
; PRIOR APPLICATION NUMBER: US/60/235,637
; PRIOR FILING DATE: 2000-09-26
; PRIOR APPLICATION NUMBER: US/60/235,638
; PRIOR FILING DATE: 2000-09-26
; PRIOR APPLICATION NUMBER: US/60/235,711
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: US/60/235,720
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: US/60/235,840
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: US/60/235,863
; PRIOR FILING DATE: 2000-09-27
; NUMBER OF SEQ ID NOS: 2276
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 1987
; LENGTH: 411
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; OTHER INFORMATION: n=a,t,g or c
US-09-954-456-1987

.....

ORGANISM: Mouse
US-10-152-661-589

Alignment Scores:

Pred. No.:	0.00103	Length:	724
Score:	98.00	Matches:	25
Percent Similarity:	43.24%	Conservative:	7
Best Local Similarity:	33.78%	Mismatches:	18
Query Match:	29.97%	Indels:	24
DB:	9	Gaps:	4

09-833799-13B (1-57) x US-10-152-661-589 (1-724)

Qy 3 GluProValLysGlyProValSerThrLysPro-----LeuIleArgCysAla 24

Db 313 AAACCACCGGAGGTCAAGTCTCCACGAAGCCACCGGCTGTGACGAGGAGGCTTAGGT 372

Qy 14 -----GlySerCysProIleIle-----LeuIleArgCysAla 24

Db 373 GTCCGAGAAAGCAGGACCTGCCCGAGCGTGGACATACCCAAAGCTCGGCTCTGT--- 429

Qy 25 MetLeuAsnProProAsnArgCysLeuLysAspThrAspCysProGlyIleLysLysCys 44

Db 430 -----GAGGACCAGTGTGACGAGGACGAGGAGGAGTGTCTGGCAACATGAATGC 477

Qy 45 CysGluGlySerCysGly---MetAlaCysPheValProGln 57

Db 478 TGCCGCAATGATGTGGGAAGATGGCCTGCACACACCCCAA 519

RESULT 40

US-09-967-768A-214

; Sequence 214, Application US/09967768A

; Patent No. US20020150877A1

; GENERAL INFORMATION:

; APPLICANT: Augustus, Meena

; TITLE OF INVENTION: Cancer Gene Determination and Therapeutic Screening Using Signatu

; FILE REFERENCE: 689290-72

; CURRENT APPLICATION NUMBER: US/09/967,768A

; CURRENT FILING DATE: 2001-09-28

; PRIOR APPLICATION NUMBER: US/60/236,109

; PRIOR FILING DATE: 2000-09-28

; PRIOR APPLICATION NUMBER: US/60/236,034

; PRIOR FILING DATE: 2000-09-28

; PRIOR APPLICATION NUMBER: US/60/236,111

; NUMBER OF SEQ ID NOS: 325

; SOFTWARE: PatentIn version 3.0

; SEQ ID NO 214

; LENGTH: 583

; TYPE: DNA

; ORGANISM: Homo sapiens

US-09-967-768A-214

Alignment Scores:

Pred. No.:	0.00103	Length:	583
Score:	97.00	Matches:	20
Percent Similarity:	53.19%	Conservative:	5
Best Local Similarity:	42.55%	Mismatches:	20
Query Match:	29.66%	Indels:	2
DB:	10	Gaps:	2

09-833799-13B (1-57) x US-09-967-768A-214 (1-583)

Qy 12 LysProGlySerCysProIleIleLeuIleArgCysAlaMetLeuAsn---ProProAsn 30

Db 256 AAGGAGGGTTCCTGCCCCCAGGTGAACATTAACCTTCCCGAGCTCGGCCTCTGTCTGGGAC 315

Qy 31 ArgCysLeuLysAspThrAspCysProGlyIleLysLysCysCysGluGlySerCysGly 50

Db 316 CAGTGCCAGGTGGACACGACGAGTGTCTTGGCCAGATGAATGCTGCCGCAATGGCTGTGGG 375

Qy 51 ---MetAlaCysPheValPro 56

Db 376 AAGGTGTCTGTGTCACTCCC 396

RESULT 41

US-09-864-761-1629

; Sequence 1629, Application US/09864761

; Patent No. US20020048763A1

; GENERAL INFORMATION:

; APPLICANT: Penn, Sharron G.

; APPLICANT: Rank, David R.

; APPLICANT: Hanzel, David K.

; APPLICANT: Chen, Wensheng

; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR

; FILE REFERENCE: Aecomica-X-1

; CURRENT APPLICATION NUMBER: US/09/864,761

; CURRENT FILING DATE: 2001-05-23

; PRIOR APPLICATION NUMBER: US 60/180,312

; PRIOR FILING DATE: 2000-02-04

; PRIOR APPLICATION NUMBER: US 60/207,456

; PRIOR FILING DATE: 2000-05-26

; PRIOR APPLICATION NUMBER: US 09/632,366

; PRIOR FILING DATE: 2000-08-03

; PRIOR APPLICATION NUMBER: GB 24263.6

; PRIOR FILING DATE: 2000-10-04

; PRIOR APPLICATION NUMBER: US 60/236,359

; PRIOR FILING DATE: 2000-09-27

; PRIOR APPLICATION NUMBER: PCT/US01/00666

; PRIOR FILING DATE: 2001-01-30

; PRIOR APPLICATION NUMBER: PCT/US01/00667

; PRIOR FILING DATE: 2001-01-30

; PRIOR APPLICATION NUMBER: PCT/US01/00664

; PRIOR FILING DATE: 2001-01-30

; PRIOR APPLICATION NUMBER: PCT/US01/00669

; PRIOR FILING DATE: 2001-01-30

; PRIOR APPLICATION NUMBER: PCT/US01/00665

; PRIOR FILING DATE: 2001-01-30

; PRIOR APPLICATION NUMBER: PCT/US01/00668

; PRIOR FILING DATE: 2001-01-30

; PRIOR APPLICATION NUMBER: PCT/US01/00663

; PRIOR FILING DATE: 2001-01-30

; PRIOR APPLICATION NUMBER: PCT/US01/00662

; PRIOR FILING DATE: 2001-01-30

; PRIOR APPLICATION NUMBER: PCT/US01/00661

; PRIOR FILING DATE: 2001-01-30

; PRIOR APPLICATION NUMBER: PCT/US01/00670

; PRIOR FILING DATE: 2001-01-30

; PRIOR APPLICATION NUMBER: US 60/234,687

; PRIOR FILING DATE: 2000-09-21

; PRIOR APPLICATION NUMBER: US 09/608,408

; PRIOR FILING DATE: 2000-06-30

; PRIOR APPLICATION NUMBER: US 09/774,203

; PRIOR FILING DATE: 2001-01-29

; NUMBER OF SEQ ID NOS: 49117

; SOFTWARE: Annomax Sequence Listing Engine vers. 1.1

; SEQ ID NO 1629

; LENGTH: 469

; TYPE: DNA

; ORGANISM: Homo sapiens

; FEATURE:

; OTHER INFORMATION: MAP TO AL050348.19

; OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 1.5

; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 1.1

; OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 1.1

; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 1.1

; OTHER INFORMATION: EXPRESSED IN BT474, SIGNAL = 1.5

US-09-864-761-1629

Alignment Scores:

Pred. No.:	0.00138	Length:	469
Score:	95.00	Matches:	19
Percent Similarity:	47.37%	Conservative:	8
Best Local Similarity:	33.33%	Mismatches:	22

```
Query Match: 29.05% Indels: 8
DB: 10 Gaps: 2

09-833799-13B (1-57) x US-09-864-761-1629 (1-469)

QY 4 ProValLysGlyProValSerThrLysPro-----GlySerCysProIleIle 19
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 228 CCTGTCCTTTTCCTGTGCTGTTGTTCCACAGGCGGGCGGTGATTGTCAAAAGTT 287
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

QY 20 LeuIleArgCysAlaMetLeuAsnProProAsnArgCysLeuLysAspThrAspCysPro 39
||||| : : : : : ||||| : : : : : ||||| : : : : : |||||
Db 288 CTGTGGGCCTGTGCATTGTTGCC-----TGTGTGATGGATGAGATTGTCAA 335
||||| : : : : : ||||| : : : : : ||||| : : : : : |||||

QY 40 GlyIleLysLysCysCysGluGlySerCysGlyMetAlaCysPheValPro 56
: : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 336 GCTGGAGAAAATGTTGCAAGTCAGGCTGTGGCCGCTTCTGTGTCCACCA 386
||||| : : : : : ||||| : : : : : ||||| : : : : : |||||

RESULT 42
US-09-790-264-57
; Sequence 57, Application US/09790264
; Patent No. US20020028508A1
; GENERAL INFORMATION:
; APPLICANT: Holtzman, Douglas A.
; APPLICANT: Goodearl, Andrew D.J.
; APPLICANT: McCarthy, Sean A.
; TITLE OF INVENTION: NOVEL GENES ENCODING PROTEINS HAVING
; TITLE OF INVENTION: PROGNASTIC, DIAGNOSTIC, PREVENTIVE, THERAPEUTIC, AND OTHER
; TITLE OF INVENTION: USES
; FILE REFERENCE: 07334-322001
; CURRENT APPLICATION NUMBER: US/09/790,264
; CURRENT FILING DATE: 2001-02-21
; PRIOR APPLICATION NUMBER: US 09/065,661
; PRIOR FILING DATE: 1998-04-23
; PRIOR APPLICATION NUMBER: US 09/298,531
; PRIOR FILING DATE: 1999-04-23
; PRIOR APPLICATION NUMBER: US 09/065,363
; PRIOR FILING DATE: 1998-04-23
; PRIOR APPLICATION NUMBER: US 09/337,930
; PRIOR FILING DATE: 1999-06-22
; PRIOR APPLICATION NUMBER: US 09/124,538
; PRIOR FILING DATE: 1998-07-29
; NUMBER OF SEQ ID NOS: 68
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 57
; LENGTH: 228
; TYPE: DNA
; ORGANISM: Mus musculus
US-09-790-264-57

Alignment Scores:
Pred. No.: 0.000616 Length: 228
Score: 94.50 Matches: 19
Percent Similarity: 41.51% Conservatve: 3
Best Local Similarity: 35.85% Mismatches: 14
Query Match: 28.90% Indels: 17
DB: 10 Gaps: 2

09-833799-13B (1-57) x US-09-790-264-57 (1-228)

QY 12 LysProGlySerCysProIleIleLeuIleArgCysAlaMetLeuAsnProPro----- 29
||||| : : : : : ||||| : : : : : ||||| : : : : : |||||
Db 91 AAACCTGGAGCTTGTCCTCC-----AAGCCTTCACCAGAA 123
||||| : : : : : ||||| : : : : : ||||| : : : : : |||||

QY 30 -----AsnArgCysLeuLysAspThrAspCysProGlyIleLysLys 43
: : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 124 AGTGTGGAATTTGTGTGATCAATGCTCAGGAGATGGATCCTGCCCTGGCAACATGAAG 183
||||| : : : : : ||||| : : : : : ||||| : : : : : |||||

QY 44 CysCysGluGlySerCysGlyMetAlaCysPheValPro 56
||||| : : : : : ||||| : : : : : ||||| : : : : : |||||
```

```
Db 184 TGCTGTAGCAATAGCTGTGGTCATGTCTGCAAAACTCCT 222

RESULT 43
US-09-823-038A-29
; Sequence 29, Application US/09823038A
; Patent No. US20020058335A1
; GENERAL INFORMATION:
; APPLICANT: Strachan, Lorna
; APPLICANT: Sleeman, Matthew
; APPLICANT: Abernethy, Nevin
; APPLICANT: Onrust, Rene
; APPLICANT: Kumble, Anand
; APPLICANT: Murison, Greg
; TITLE OF INVENTION: Compositions Isolated From Stromal Cells
; TITLE OF INVENTION: and Methods For Their Use
; FILE REFERENCE: 11000.1037c3
; CURRENT APPLICATION NUMBER: US/09/823,038A
; CURRENT FILING DATE: 2001-07-09
; NUMBER OF SEQ ID NOS: 61
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 29
; LENGTH: 437
; TYPE: DNA
; ORGANISM: Mouse
US-09-823-038A-29

Alignment Scores:
Pred. No.: 0.00146 Length: 437
Score: 94.50 Matches: 19
Percent Similarity: 41.51% Conservatve: 3
Best Local Similarity: 35.85% Mismatches: 14
Query Match: 28.90% Indels: 17
DB: 10 Gaps: 2

09-833799-13B (1-57) x US-09-823-038A-29 (1-437)

QY 12 LysProGlySerCysProIleIleLeuIleArgCysAlaMetLeuAsnProPro----- 29
||||| : : : : : ||||| : : : : : ||||| : : : : : |||||
Db 80 AAACCTGGAGCTTGTCCTCC-----AAGCCTTCACCAGAA 112
||||| : : : : : ||||| : : : : : ||||| : : : : : |||||

QY 30 -----AsnArgCysLeuLysAspThrAspCysProGlyIleLysLys 43
: : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 113 AGTGTGGAATTTGTGTGATCAATGCTCAGGAGATGGATCCTGCCCTGGCAACATGAAG 172
||||| : : : : : ||||| : : : : : ||||| : : : : : |||||

QY 44 CysCysGluGlySerCysGlyMetAlaCysPheValPro 56
||||| : : : : : ||||| : : : : : ||||| : : : : : |||||
Db 173 TGCTGTAGCAATAGCTGTGGTCATGTCTGCAAAACTCCT 211
||||| : : : : : ||||| : : : : : ||||| : : : : : |||||

RESULT 44
US-09-790-264-55
; Sequence 55, Application US/09790264
; Patent No. US20020028508A1
; GENERAL INFORMATION:
; APPLICANT: Holtzman, Douglas A.
; APPLICANT: Goodearl, Andrew D.J.
; APPLICANT: McCarthy, Sean A.
; TITLE OF INVENTION: NOVEL GENES ENCODING PROTEINS HAVING
; TITLE OF INVENTION: PROGNASTIC, DIAGNOSTIC, PREVENTIVE, THERAPEUTIC, AND OTHER
; TITLE OF INVENTION: USES
; FILE REFERENCE: 07334-322001
; CURRENT APPLICATION NUMBER: US/09/790,264
; CURRENT FILING DATE: 2001-02-21
; PRIOR APPLICATION NUMBER: US 09/065,661
; PRIOR FILING DATE: 1998-04-23
; PRIOR APPLICATION NUMBER: US 09/298,531
; PRIOR FILING DATE: 1999-04-23
; PRIOR APPLICATION NUMBER: US 09/065,363
; PRIOR FILING DATE: 1998-04-23
; PRIOR APPLICATION NUMBER: US 09/337,930
; PRIOR FILING DATE: 1999-06-22
; PRIOR APPLICATION NUMBER: US 09/102,705
; PRIOR FILING DATE: 1998-06-22
; PRIOR APPLICATION NUMBER: US 09/363,630
```

```

; PRIOR FILING DATE: 1999-07-29
; PRIOR APPLICATION NUMBER: US 09/124,538
; PRIOR FILING DATE: 1998-07-29
; NUMBER OF SEQ ID NOS: 68
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 55
; LENGTH: 471
; TYPE: DNA
; ORGANISM: Mus musculus
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (37) ... (264)
US-09-790-264-55

Alignment Scores:
Pred. No.: 0.00161 Length: 471
Score: 94.50 Matches: 19
Percent Similarity: 41.51% Conservative: 3
Best Local Similarity: 35.85% Mismatches: 14
Query Match: 28.90% Indels: 17
DB: 10 Gaps: 2

09-833799-13B (1-57) x US-09-790-264-55 (1-471)

QY 12 LysProGlySerCysProIleLeuIleArgCysAlaMetLeuAsnProPro----- 29
Db 127 AAACCTGGAGCTTGCTCC-----AAGCTTCACCAGAA 159

QY 30 -----AsnArgCysLeuLysAspThrAspCysProGlyIleLysLys 43
Db 160 AGTGTGGAATTGTGTGATCAATGCTCAGGAGATGGATCTGCCCTGGCAACATGAAG 219

QY 44 CysCysGluGlySerCysGlyMetAlaCysPheValPro 56
Db 220 TGCTGTAGCAATAGCTGTGTCATGTCTGCAAAACTCCT 258

RESULT 45
US-09-925-300-421
; Sequence 421, Application US/09925300
; Patent No. US20020151681A1
; GENERAL INFORMATION:
; APPLICANT: Craig Rosen,
; APPLICANT: Steve Ruben
; TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies
; FILE REFERENCE: PA101
; CURRENT APPLICATION NUMBER: US/09/925,300
; CURRENT FILING DATE: 2001-08-10
; PRIOR APPLICATION NUMBER: PCT/US00/05988
; PRIOR FILING DATE: 2000-03-08
; PRIOR APPLICATION NUMBER: 60/124,270
; PRIOR FILING DATE: 1999-03-12
; NUMBER OF SEQ ID NOS: 1890
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 421
; LENGTH: 622
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-925-300-421

Alignment Scores:
Pred. No.: 0.00359 Length: 622
Score: 93.00 Matches: 19
Percent Similarity: 53.19% Conservative: 6
Best Local Similarity: 40.43% Mismatches: 20
Query Match: 28.44% Indels: 2
DB: 10 Gaps: 2

09-833799-13B (1-57) x US-09-925-300-421 (1-622)

QY 12 LysProGlySerCysProIleLeuIleArgCysAlaMetLeuAsn---ProProAsn 30
Db 288 AAGGAGGGTTCCTGCTGCCCCCAGGTGAACATTAACTTTCCCGAGCTCGGCCTCTGTGCGGAC 347
```

```

QY 31 ArgCysLeuLysAspThrAspCysProGlyIleLysLysCysCysGluGlySerCysGly 50
Db 348 CAGTGCCAGGTGGACAGCCAGTGTCTCTGCGCCAGATGAAATGCTGCCGCAATGGCTGTGGG 407

QY 51 ---MetAlaCysPheValPro 56
Db 408 AAGGTGCTCTGTGTCACTCCC 428

RESULT 46
US-09-924-340-41
; Sequence 41, Application US/09924340
; Publication No. US20030027248A1
; GENERAL INFORMATION:
; APPLICANT: Bejanin, Stephanie
; APPLICANT: Tanaka, Hiroaki
; TITLE OF INVENTION: HUMAN CDNAS AND PROTEINS AND USES THEREOF
; FILE REFERENCE: 91.US2.REG
; CURRENT APPLICATION NUMBER: US/09/924,340
; CURRENT FILING DATE: 2001-08-06
; PRIOR APPLICATION NUMBER: US 60/305,456
; PRIOR FILING DATE: 2001-07-13
; PRIOR APPLICATION NUMBER: US 60/302,277
; PRIOR FILING DATE: 2001-06-29
; PRIOR APPLICATION NUMBER: US 60/298,698
; PRIOR FILING DATE: 2001-06-15
; PRIOR APPLICATION NUMBER: US 60/293,574
; PRIOR FILING DATE: 2001-05-25
; NUMBER OF SEQ ID NOS: 112
; SOFTWARE: JPatent
; SEQ ID NO 41
; LENGTH: 643
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: 5'UTR
; LOCATION: 1..90
; NAME/KEY: CDS
; LOCATION: 91..462
; NAME/KEY: 3'UTR
; LOCATION: 463..643
; NAME/KEY: polyA signal
; LOCATION: 607..612
; NAME/KEY: polyA site
; LOCATION: 628..643
US-09-924-340-41

Alignment Scores:
Pred. No.: 0.00375 Length: 643
Score: 93.00 Matches: 19
Percent Similarity: 53.19% Conservative: 6
Best Local Similarity: 40.43% Mismatches: 20
Query Match: 28.44% Indels: 2
DB: 9 Gaps: 2

09-833799-13B (1-57) x US-09-924-340-41 (1-643)

QY 12 LysProGlySerCysProIleLeuIleArgCysAlaMetLeuAsn---ProProAsn 30
Db 316 AAGGAGGGTTCCTGCTGCCCCCAGGTGAACATTAACTTTCCCGAGCTCGGCCTCTGTGCGGAC 375

QY 31 ArgCysLeuLysAspThrAspCysProGlyIleLysLysCysCysGluGlySerCysGly 50
Db 376 CAGTGCCAGGTGGACAGCCAGTGTCTCTGCGCCAGATGAAATGCTGCCGCAATGGCTGTGGG 435

QY 51 ---MetAlaCysPheValPro 56
Db 436 AAGGTGCTCTGTGTCACTCCC 456

RESULT 47
US-09-992-600A-41
; Sequence 41, Application US/09992600A
; Publication No. US20030027161A1
; GENERAL INFORMATION:
```


Db 538 AAGCAGGGGACTGTCGGGCTCTGAGAAAGCCAGTGGATTGGGGCCGCTGTGT--- 594
Qy 26 LeuAsnProProAsnArgCysLeuLysAspThrAspCysProGlyIleLeuLysCysCys 45
Db 595 -----GAAAGCTGCGAAGTTGACAATGAGTGTCTCTGGGGTGAAGAAATGTTGT 642
Qy 46 GluGlySerCysGlyMetAlaCysPheValProGln 57
Db 643 TCGAATGGGTGTGGACACACCTGTCAAGTACCCCAAG 678

RESULT 51
US-09-852-797-34
; Sequence 34, Application US/09852797
; Patent No. US20020172994A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: 28 Human Secreted Proteins
; FILE REFERENCE: PZ003P2
; CURRENT APPLICATION NUMBER: US/09/852,797
; CURRENT FILING DATE: 2001-05-11
; PRIOR APPLICATION NUMBER: 60/265,583
; PRIOR FILING DATE: 2001-02-02
; PRIOR APPLICATION NUMBER: 09/152,060
; PRIOR FILING DATE: 1998-09-11
; PRIOR APPLICATION NUMBER: PCT/US98/04858
; PRIOR FILING DATE: 1998-03-12
; PRIOR APPLICATION NUMBER: 60/040,762
; PRIOR FILING DATE: 1997-03-14
; PRIOR APPLICATION NUMBER: 60/040,710
; PRIOR FILING DATE: 1997-03-14
; PRIOR APPLICATION NUMBER: 60/050,934
; PRIOR FILING DATE: 1997-05-30
; PRIOR APPLICATION NUMBER: 60/048,100
; PRIOR FILING DATE: 1997-05-30
; PRIOR APPLICATION NUMBER: 60/048,357
; PRIOR FILING DATE: 1997-05-30
; PRIOR APPLICATION NUMBER: 60/048,189
; PRIOR FILING DATE: 1997-05-30
; PRIOR APPLICATION NUMBER: 60/057,765
; PRIOR FILING DATE: 1997-09-05
; PRIOR APPLICATION NUMBER: 60/048,970
; PRIOR FILING DATE: 1997-06-06
; PRIOR APPLICATION NUMBER: 60/068,368
; PRIOR FILING DATE: 1997-12-19
; NUMBER OF SEQ ID NOS: 118
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 34
; LENGTH: 753
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-852-797-34

Alignment Scores:
Pred. No.: 0.0471 Length: 753
Score: 85.00 Matches: 21
Percent Similarity: 44.23% Conservative: 2
Best Local Similarity: 40.38% Mismatches: 25
Query Match: 25.99% Indels: 4
DB: 9 Gaps: 2

09-833799-13B (1-57) x US-09-852-797-34 (1-753)

Qy 5 VallysGlyProValSerThrLys-----ProGlySerCysProIleIleLeu 21
Db 85 GTCCAGGGACCTGGTCTGACTGATTGGTTATTTCCCGAGAGATGTCCCAAAATCAGAGAA 144
Qy 22 ArgCysAlaMetLeuAsnProProAsnArgCysLeuLysAspThrAspCysProGlyIle 41
Db 145 GAATGTGAA---TTCCAAGAAAGGATGTGTGTACAAAGCAGACACAATGCCAGGACAAC 201
Qy 42 LysLysCysCysGluGlySerCysGlyMetAlaCys 53

Alignment Scores:
Pred. No.: 0.0891 Length: 6314
Score: 92.50 Matches: 18
Percent Similarity: 40.38% Conservative: 3
Best Local Similarity: 34.62% Mismatches: 20
Query Match: 28.29% Indels: 11
DB: 10 Gaps: 2

09-833799-13B (1-57) x US-09-954-456-319 (1-6314)

Qy 12 LysProGlySerCysProIle-----IleLeuIleArgCysAlaMet 25
Db 538 AAGCAGGGGACTGTCGGGCTCTGAGAAAGCCAGTGGATTGGGGCCGCTGTGT--- 594
Qy 26 LeuAsnProProAsnArgCysLeuLysAspThrAspCysProGlyIleLeuLysCysCys 45
Db 595 -----GAAAGCTGCGAAGTTGACAATGAGTGTCTCTGGGGTGAAGAAATGTTGT 642
Qy 46 GluGlySerCysGlyMetAlaCysPheValProGln 57
Db 643 TCGAATGGGTGTGGACACACCTGTCAAGTACCCCAAG 678

RESULT 50
US-09-954-456-1230
; Sequence 1230, Application US/09954456
; Patent No. US20020115057A1
; GENERAL INFORMATION:
; APPLICANT: Young, Paul
; TITLE OF INVENTION: Process for Identifying Anti-Cancer Therapeutic Agents Using Cand
; TITLE OF INVENTION: Sets
; FILE REFERENCE: 689290-76
; CURRENT APPLICATION NUMBER: US/09/954,456
; CURRENT FILING DATE: 2001-09-18
; PRIOR APPLICATION NUMBER: US/60/233,617
; PRIOR FILING DATE: 2000-09-18
; PRIOR APPLICATION NUMBER: US/60/234,052
; PRIOR FILING DATE: 2000-09-20
; PRIOR APPLICATION NUMBER: US/60/234,923
; PRIOR FILING DATE: 2000-09-25
; PRIOR APPLICATION NUMBER: US/60/235,134
; PRIOR FILING DATE: 2000-09-25
; PRIOR APPLICATION NUMBER: US/60/235,637
; PRIOR FILING DATE: 2000-09-26
; PRIOR APPLICATION NUMBER: US/60/235,638
; PRIOR FILING DATE: 2000-09-26
; PRIOR APPLICATION NUMBER: US/60/235,711
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: US/60/235,720
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: US/60/235,840
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: US/60/235,863
; PRIOR FILING DATE: 2000-09-27
; NUMBER OF SEQ ID NOS: 2276
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 1230
; LENGTH: 6314
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-954-456-1230

Alignment Scores:
Pred. No.: 0.0891 Length: 6314
Score: 92.50 Matches: 18
Percent Similarity: 40.38% Conservative: 3
Best Local Similarity: 34.62% Mismatches: 20
Query Match: 28.29% Indels: 11
DB: 10 Gaps: 2

09-833799-13B (1-57) x US-09-954-456-1230 (1-6314)

Qy 12 LysProGlySerCysProIle-----IleLeuIleArgCysAlaMet 25

Db 202 AAGAAGTGTGTGTTCTTTCAGCTGCGGAAAAAAATGT 237

RESULT 52

US-09-853-161-34

; Sequence 34, Application US/09853161

; Patent No. US20020076756A1

; GENERAL INFORMATION:

; APPLICANT: Rosen et al.

; TITLE OF INVENTION: 28 Human Secreted Proteins

; FILE REFERENCE: PZ003P3

; CURRENT APPLICATION NUMBER: US/09/853,161

; CURRENT FILING DATE: 2001-05-11

; PRIOR APPLICATION NUMBER: 60/265,583

; PRIOR FILING DATE: 2001-02-02

; PRIOR APPLICATION NUMBER: 09/152,060

; PRIOR FILING DATE: 1998-09-11

; PRIOR APPLICATION NUMBER: PCT/US98/04858

; PRIOR FILING DATE: 1998-03-12

; PRIOR APPLICATION NUMBER: 60/040,762

; PRIOR FILING DATE: 1997-03-14

; PRIOR APPLICATION NUMBER: 60/040,710

; PRIOR FILING DATE: 1997-03-14

; PRIOR APPLICATION NUMBER: 60/050,934

; PRIOR FILING DATE: 1997-05-30

; PRIOR APPLICATION NUMBER: 60/048,100

; PRIOR FILING DATE: 1997-05-30

; PRIOR APPLICATION NUMBER: 60/048,357

; PRIOR FILING DATE: 1997-05-30

; PRIOR APPLICATION NUMBER: 60/048,189

; PRIOR FILING DATE: 1997-05-30

; PRIOR APPLICATION NUMBER: 60/057,765

; PRIOR FILING DATE: 1997-09-05

; PRIOR APPLICATION NUMBER: 60/048,970

; PRIOR FILING DATE: 1997-06-06

; PRIOR APPLICATION NUMBER: 60/068,368

; PRIOR FILING DATE: 1997-12-19

; NUMBER OF SEQ ID NOS: 118

; SOFTWARE: PatentIn Ver. 2.0

; SEQ ID NO 34

; LENGTH: 753

; TYPE: DNA

; ORGANISM: Homo sapiens

US-09-853-161-34

Alignment Scores:

Pred. No.:	0.0471	Length:	753
Score:	85.00	Matches:	21
Percent Similarity:	44.23%	Conservative:	2
Best Local Similarity:	40.38%	Mismatches:	25
Query Match:	25.99%	Indels:	4
DB:	10	Gaps:	2

09-833799-13B (1-57) x US-09-853-161-34 (1-753)

QY 5 ValLysGlyProValSerThrLys-----ProGlySerCysProIleIleLeuIle 21

Db 85 GTCCAGGGACCTGGTCTGACTGATTGGTTATTTCCTCCAGGAGATGTCCCAAAATCAGAGAA 144

QY 22 ArgCysAlaMetLeuAsnProProAsnArgCysLeuLysAspThrAspCysProGlyIle 41

Db 145 GAATGTGAA---TTCCAAGAAAGGATGTGTGTACAAAGGACAGACAATGCCAGGACAAC 201

QY 42 LysLysCysCysGluGlySerCysGlyMetAlaCys 53

Db 202 AAGAAGTGTGTGTTCTTTCAGCTGCGGAAAAAAATGT 237

RESULT 53

US-09-852-659A-34

; Sequence 34, Application US/09852659A

; Patent No. US20020077287A1

; GENERAL INFORMATION:

; APPLICANT: Rosen et al.

; TITLE OF INVENTION: 28 Human Secreted Proteins

; FILE REFERENCE: PZ003P4

; CURRENT APPLICATION NUMBER: US/09/852,659A

; CURRENT FILING DATE: 2001-05-11

; PRIOR APPLICATION NUMBER: 60/265,583

; PRIOR FILING DATE: 2001-02-02

; PRIOR APPLICATION NUMBER: 09/152,060

; PRIOR FILING DATE: 1998-09-11

; PRIOR APPLICATION NUMBER: PCT/US98/04858

; PRIOR FILING DATE: 1998-03-12

; PRIOR APPLICATION NUMBER: 60/040,762

; PRIOR FILING DATE: 1997-03-14

; PRIOR APPLICATION NUMBER: 60/040,710

; PRIOR FILING DATE: 1997-03-14

; PRIOR APPLICATION NUMBER: 60/050,934

; PRIOR FILING DATE: 1997-05-30

; PRIOR APPLICATION NUMBER: 60/048,100

; PRIOR FILING DATE: 1997-05-30

; PRIOR APPLICATION NUMBER: 60/048,357

; PRIOR FILING DATE: 1997-05-30

; PRIOR APPLICATION NUMBER: 60/048,189

; PRIOR FILING DATE: 1997-05-30

; PRIOR APPLICATION NUMBER: 60/057,765

; PRIOR FILING DATE: 1997-09-05

; PRIOR APPLICATION NUMBER: 60/048,970

; PRIOR FILING DATE: 1997-06-06

; PRIOR APPLICATION NUMBER: 60/068,368

; PRIOR FILING DATE: 1997-12-19

; NUMBER OF SEQ ID NOS: 121

; SOFTWARE: PatentIn Ver. 2.0

; SEQ ID NO 34

; LENGTH: 753

; TYPE: DNA

; ORGANISM: Homo sapiens

US-09-852-659A-34

Alignment Scores:

Pred. No.:	0.0471	Length:	753
Score:	85.00	Matches:	21
Percent Similarity:	44.23%	Conservative:	2
Best Local Similarity:	40.38%	Mismatches:	25
Query Match:	25.99%	Indels:	4
DB:	10	Gaps:	2

09-833799-13B (1-57) x US-09-852-659A-34 (1-753)

QY 5 ValLysGlyProValSerThrLys-----ProGlySerCysProIleIleLeuIle 21

Db 85 GTCCAGGGACCTGGTCTGACTGATTGGTTATTTCCTCCAGGAGATGTCCCAAAATCAGAGAA 144

QY 22 ArgCysAlaMetLeuAsnProProAsnArgCysLeuLysAspThrAspCysProGlyIle 41

Db 145 GAATGTGAA---TTCCAAGAAAGGATGTGTGTACAAAGGACAGACAATGCCAGGACAAC 201

QY 42 LysLysCysCysGluGlySerCysGlyMetAlaCys 53

Db 202 AAGAAGTGTGTGTTCTTTCAGCTGCGGAAAAAAATGT 237

RESULT 54

US-09-852-797-50

; Sequence 50, Application US/09852797

; Patent No. US20020172994A1

; GENERAL INFORMATION:

; APPLICANT: Rosen et al.

; TITLE OF INVENTION: 28 Human Secreted Proteins

; FILE REFERENCE: PZ003P2

; CURRENT APPLICATION NUMBER: US/09/852,797

; CURRENT FILING DATE: 2001-05-11

; PRIOR APPLICATION NUMBER: 60/265,583

; PRIOR FILING DATE: 2001-02-02

; PRIOR APPLICATION NUMBER: 09/152,060

; PRIOR FILING DATE: 1998-09-11

; PRIOR APPLICATION NUMBER: PCT/US98/04858

; PRIOR FILING DATE: 1998-03-12

; PRIOR FILING DATE: 1997-05-30
 ; PRIOR APPLICATION NUMBER: 60/048,357

;; PRIOR APPLICATION NUMBER: 60/068,368
;; PRIOR FILING DATE: 1997-12-19

; NUMBER OF SEQ ID NOS: 121
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 50
; LENGTH: 783
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-852-659A-50

Alignment Scores:

Pred. No.: 0.0496 Length: 783
Score: 85.00 Matches: 21
Percent Similarity: 44.23% Conservative: 2
Best Local Similarity: 40.38% Mismatches: 25
Query Match: 25.99% Indels: 4
DB: 10 Gaps: 2

09-833799-13B (1-57) x US-09-852-659A-50 (1-783)

QY 5 ValLysGlyProValSerThrLys-----ProGlySerCysProIleIleLeuIle 21
Db 85 GTCCAGGGACCTGGTCTGACTGATTGTTATTTCCAGGAGATGTCCCAAAATCAGAGAA 144
QY 22 ArgCysAlaMetLeuAsnProProAsnArgCysLeuLysAspThrAspCysProGlyIle 41
Db 145 GAATGTGAA---TTCCAAGAAAGGATGTGTGTACAAAGGACAGACAATGCCAGGACAAC 201
QY 42 LysLysCysCysGluGlySerCysGlyMetAlaCys 53
Db 202 AAGAAGTGTGTGTCTTCAGCTGCGGAAAAAATGT 237

RESULT 57

US-09-852-659A-121
; Sequence 121, Application US/09852659A
; Patent No. US20020077287A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: 28 Human Secreted Proteins
; FILE REFERENCE: PZ003P4
; CURRENT FILING DATE: 2001-05-11
; PRIOR APPLICATION NUMBER: US/09/852,659A
; PRIOR FILING DATE: 2001-02-02
; PRIOR APPLICATION NUMBER: 60/265,583
; PRIOR FILING DATE: 1998-09-11
; PRIOR APPLICATION NUMBER: 09/152,060
; PRIOR FILING DATE: 1998-03-12
; PRIOR APPLICATION NUMBER: PCT/US98/04858
; PRIOR FILING DATE: 1997-03-14
; PRIOR APPLICATION NUMBER: 60/040,762
; PRIOR FILING DATE: 1997-03-14
; PRIOR APPLICATION NUMBER: 60/040,710
; PRIOR FILING DATE: 1997-03-14
; PRIOR APPLICATION NUMBER: 60/050,934
; PRIOR FILING DATE: 1997-05-30
; PRIOR APPLICATION NUMBER: 60/048,100
; PRIOR FILING DATE: 1997-05-30
; PRIOR APPLICATION NUMBER: 60/048,357
; PRIOR FILING DATE: 1997-05-30
; PRIOR APPLICATION NUMBER: 60/048,189
; PRIOR FILING DATE: 1997-05-30
; PRIOR APPLICATION NUMBER: 60/057,765
; PRIOR FILING DATE: 1997-09-05
; PRIOR APPLICATION NUMBER: 60/048,970
; PRIOR FILING DATE: 1997-06-06
; PRIOR APPLICATION NUMBER: 60/068,368
; PRIOR FILING DATE: 1997-12-19
; NUMBER OF SEQ ID NOS: 121
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 121
; LENGTH: 354
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-852-659A-121

Alignment Scores:

Pred. No.: 0.064 Length: 354
Score: 80.50 Matches: 17
Percent Similarity: 45.00% Conservative: 1
Best Local Similarity: 42.50% Mismatches: 21
Query Match: 24.62% Indels: 1
DB: 10 Gaps: 1

09-833799-13B (1-57) x US-09-852-659A-121 (1-354)

QY 14 GlySerCysProIleIleLeuIleArgCysAlaMetLeuAsnProProAsnArgCysLeu 33
Db 43 GGGAGATGTCCCAAAATCAGAGAAGAAATGTGAA---TTCCAAGAAAGGATGTGTGTACA 99
QY 34 LysAspThrAspCysProGlyIleLysLysCysCysGluGlySerCysGlyMetAlaCys 53
Db 100 AAGACAGACAATGCCAGACAACAAGAGTGTGTCTTCAGCTGCGGAAAAAATGT 159

RESULT 58

US-09-833-381-747/c
; Sequence 747, Application US/09833381
; Patent No. US20020132090A1
; GENERAL INFORMATION:
; APPLICANT: Robison, Keith E.
; TITLE OF INVENTION: No. US20020132090A1el Nucleic Acid and Protein Homologs
; FILE REFERENCE: 5800-119
; CURRENT APPLICATION NUMBER: US/09/833,381
; CURRENT FILING DATE: 2001-04-11
; PRIOR APPLICATION NUMBER: 09/516,448
; PRIOR FILING DATE: 2000-02-29
; NUMBER OF SEQ ID NOS: 2050
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 747
; LENGTH: 334
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(334)
; OTHER INFORMATION: n = A,T,C or G
US-09-833-381-747

Alignment Scores:

Pred. No.: 0.106 Length: 334
Score: 78.50 Matches: 20
Percent Similarity: 45.16% Conservative: 8
Best Local Similarity: 32.26% Mismatches: 21
Query Match: 24.01% Indels: 13
DB: 10 Gaps: 2

09-833799-13B (1-57) x US-09-833-381-747 (1-334)

QY 2 GlnGluProValLysGlyProValSerThrLysProGlySerCysProIleIleLeuIle 21
Db 300 AAGAAGCCATTGGCGGGGCCCATTTCCGCCAAGGAAGGCTCATCGGGGAAGGAGACCCCTC 241
QY 22 ArgCysAlaMetLeuAsnProProAsnArgCysLeuLysAspThrAspCysProGly--- 40
Db 240 AGCGTCGCTGTCTGCCCTCCCTGAACGTGT-----CCTTGTCGACTTGG 193
QY 41 -----IleLysLysCysCysGluGlySerCysGlyMetAla 52
Db 192 AGCCGCGCGGGAGCGTTCATGTTCGGGAAATACTGCCCCCAGCGCTGCGGGCGCGG 133
QY 53 CysPhe 54
Db 132 TCCTTC 127

RESULT 59

US-09-819-136-1
; Sequence 1, Application US/09819136
; Patent No. US20020146789A1
; GENERAL INFORMATION:
; APPLICANT: Conklin, Darrell C.

; APPLICANT: Gao, Zeren
; TITLE OF INVENTION: MULTI-DOMAIN PROTEINASE INHIBITOR
; FILE REFERENCE: 00-25
; CURRENT APPLICATION NUMBER: US/09/819,136
; CURRENT FILING DATE: 2001-03-27
; PRIOR APPLICATION NUMBER: US 60/193,642
; PRIOR FILING DATE: 2000-03-31
; NUMBER OF SEQ ID NOS: 13
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 1
; LENGTH: 2082
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (376)...(2022)
US-09-819-136-1

Alignment Scores:
Pred. No.: 3.8 Length: 2082
Score: 74.50 Matches: 24
Percent Similarity: 26.26% Conservative: 9
Best Local Similarity: 36.37% Mismatches: 16
Query Match: 22.78% Indels: 43
DB: 10 Gaps: 4

09-833799-13B (1-57) x US-09-819-136-1 (1-2082)
Qy 3 GluProValLysGlyProValSerThrLysProGly-----SerCys 16
Db 320 GAGCCCGAGGAGGGCTGGTGGCCACACCCCGCGCTGGCTCGGCGGCGCTCATGC 379
Qy 17 ProIleIleLeuIleArgCysAlaMetLeu----- 26
Db 380 CCGCCCTAGCTCCACTCTGCGG-CTCTTGCTCTCTCGGCTGACCTCGGGGGCTGGC 438
Qy 27 -----AsnPro-----ProAsnArg----- 31
Db 439 TTGCTGCCAGGGCTGGGAGCCACCGCGGCTGTGCCCCAACACAGCTCAGCCCAACCTG 498
Qy 32 -----CysLeuLysAspThrAspCysProGly 40
Db 499 TGGGTGGAGCGCCAGAGCACCTGTAGCGCGAGTGTAGCAGGGACCAAGACTGTGCGGCT 558
Qy 41 IleLysLysCysCysGluGlySerCysGlyMet 51
Db 559 GCTGAGAGTGTGCTGCATCAACAGTGTGTGGACTG 591

RESULT 60
US-09-894-882-240
; Sequence 240, Application US/09894882
; Patent No. US20020102607A1
; GENERAL INFORMATION:
; APPLICANT: University of Utah Research Foundation
; APPLICANT: Cognetix, Inc.
; APPLICANT: Walker, Craig S.
; APPLICANT: Shetty, Reshma
; APPLICANT: Jimenez, Elsie C.
; APPLICANT: McIntosh, J. Michael
; APPLICANT: Olivera, Baldomero M.
; APPLICANT: Watkins, Maren
; APPLICANT: Jones, Robert M.
; APPLICANT: Shen, Greg S.
; TITLE OF INVENTION: I-Superfamily Conotoxins
; FILE REFERENCE: 2314-238
; CURRENT APPLICATION NUMBER: US/09/894,882
; CURRENT FILING DATE: 2001-06-29
; PRIOR APPLICATION NUMBER: US 60/
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: US 60/243,410
; PRIOR FILING DATE: 2000-10-27
; PRIOR APPLICATION NUMBER: US 60/246,581
; PRIOR FILING DATE: 2000-11-08

; PRIOR APPLICATION NUMBER: US 60/247,714
; PRIOR FILING DATE: 2000-11-14
; PRIOR APPLICATION NUMBER: US 60/264,256
; PRIOR FILING DATE: 2001-01-29
; NUMBER OF SEQ ID NOS: 506
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 240
; LENGTH: 243
; TYPE: DNA
; ORGANISM: Conus emaciatus
US-09-894-882-240
Alignment Scores:
Pred. No.: 0.296 Length: 243
Score: 73.50 Matches: 18
Percent Similarity: 54.17% Conservative: 8
Best Local Similarity: 37.50% Mismatches: 15
Query Match: 22.48% Indels: 7
DB: 10 Gaps: 3
09-833799-13B (1-57) x US-09-894-882-240 (1-243)
Qy 15 SerCysProIleIleLeuIleArgCysAla---MetLeuAsnProProAsnArgCysLeu 33
Db 25 AGCTGTTTCTACTACTGTCATCGTTTGTCTGAACCTGATTGTGCTTATCAATGCTGCTAC 84
Qy 34 LysAsp---ThrAspCysProGlyIle-----LysLysCysCysGluGly 47
Db 85 CAAGATGAAACGCCCTGCAGAGGTAGTATCTTCTGCTGTGCGCAAAAATGCTGTATAGGG 144
Qy 48 SerCysGlyMetAlaCysPheVal 55
Db 145 ACATGCAGATTTCGGTGTACGTT 168

RESULT 61
US-09-894-882-249
; Sequence 249, Application US/09894882
; Patent No. US20020102607A1
; GENERAL INFORMATION:
; APPLICANT: University of Utah Research Foundation
; APPLICANT: Cognetix, Inc.
; APPLICANT: Walker, Craig S.
; APPLICANT: Shetty, Reshma
; APPLICANT: Jimenez, Elsie C.
; APPLICANT: McIntosh, J. Michael
; APPLICANT: Olivera, Baldomero M.
; APPLICANT: Watkins, Maren
; APPLICANT: Jones, Robert M.
; APPLICANT: Shen, Greg S.
; TITLE OF INVENTION: I-Superfamily Conotoxins
; FILE REFERENCE: 2314-238
; CURRENT APPLICATION NUMBER: US/09/894,882
; CURRENT FILING DATE: 2001-06-29
; PRIOR APPLICATION NUMBER: US 60/
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: US 60/243,410
; PRIOR FILING DATE: 2000-10-27
; PRIOR APPLICATION NUMBER: US 60/246,581
; PRIOR FILING DATE: 2000-11-08
; PRIOR APPLICATION NUMBER: US 60/247,714
; PRIOR FILING DATE: 2000-11-14
; PRIOR APPLICATION NUMBER: US 60/264,256
; PRIOR FILING DATE: 2001-01-29
; NUMBER OF SEQ ID NOS: 506
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 249
; LENGTH: 243
; TYPE: DNA
; ORGANISM: Conus emaciatus
US-09-894-882-249
Alignment Scores:
Pred. No.: 0.296 Length: 243

;; PRIOR FILING DATE: 2000-09-25
;; PRIOR APPLICATION NUMBER: US/60/235,280
;; PRIOR FILING DATE: 2000-09-25
;; NUMBER OF SEQ ID NOS: 259
;; SOFTWARE: PatentIn version 3.0
;; SEQ ID NO 225
;; LENGTH: 2108
;; TYPE: DNA
;; ORGANISM: Homo sapiens
US-09-962-832-225

Alignment Scores:
Pred. No.: 29.4 Length: 2108
Score: 67.50 Matches: 18
Percent Similarity: 46.67% Conservative: 3
Best Local Similarity: 40.00% Mismatches: 17
Query Match: 20.64% Indels: 7
DB: 10 Gaps: 3

09-833799-13B (1-57) x US-09-962-832-225 (1-2108)

Qy 10 SerThrLysProGlySerCysProIleLeuIleArgCysAlaMetLeuAsnProPro 29
Db 1042 TCCTCCAGCTGCTCCAGTTGCCCTCCTGCTGCTCCAGATGC-----TTAGGCTGCCCC 989
Qy 30 AsnArgCysLeuLysAspThrAspCysProGlyIleLysLysCysCysGluGlySerCys 49
Db 988 TCCTGCTGCTCCAGGTGCTTACGTGCCCC-----ATCTGCTGCTCTGGGAGCTCT 938
Qy 50 GlyMetAlaCysPhe 54
Db 937 GGC-----TGCTTC 929

RESULT 70
US-09-962-832-154
; Sequence 154, Application US/09962832
; Patent No. US20020110821A1
; GENERAL INFORMATION:
; APPLICANT: Ebner, Reinhard
; TITLE OF INVENTION: Cancer Gene Determination and Therapeutic Screening Using Signatu
; FILE REFERENCE: 689290-74
; CURRENT APPLICATION NUMBER: US/09/962,832
; CURRENT FILING DATE: 2001-09-25
; PRIOR APPLICATION NUMBER: US/60/235,077
; PRIOR FILING DATE: 2000-09-25
; PRIOR APPLICATION NUMBER: US/60/235,280
; PRIOR FILING DATE: 2000-09-25
; NUMBER OF SEQ ID NOS: 259
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 154
; LENGTH: 302250
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-962-832-154

Alignment Scores:
Pred. No.: 2.39e+04 Length: 302250
Score: 67.00 Matches: 22
Percent Similarity: 43.33% Conservative: 4
Best Local Similarity: 36.67% Mismatches: 19
Query Match: 20.49% Indels: 15
DB: 10 Gaps: 3

09-833799-13B (1-57) x US-09-962-832-154 (1-302250)

Qy 4 ProValLysGlyProValSerThrLysProGlySerCysProIleLeuIleLeuLeuArgCys 23
Db 291823 CCAGTCAAGTTCCTGTTAGTGA-----AGCTGCCCTTCCCTTGTTCATAATGT 291873
Qy 24 AlaMetLeuAsnProProAsnArgCysLeu-LysAspThrAspCysProGlyIleLysLy 43
Db 291874 CCCTGC-----TGCTGCCAGGGGAAGGAGTGTCTTCGTTCTCTGAG 291915

Qy 43 sCysCysGluGlySerCysGlyMetAla-----CysPheValProGln 57
Db 291916 CCTATGTTTCCAGGGCTTGGGATGGCTCATGGCCCCAGCCCGCTGCTCGTTTCCCAA 291973
RESULT 71
US-09-815-242-7863/c
; Sequence 7863, Application US/09815242
; Patent No. US20020061569A1
; GENERAL INFORMATION:
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari L.
; APPLICANT: Zyskind, Judith W.
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John D.
; APPLICANT: Carr, Grant J.
; APPLICANT: Yamamoto, Robert T.
; APPLICANT: Xu, H. Howard
; TITLE OF INVENTION: Identification of Essential Genes in
; TITLE OF INVENTION: Prokaryotes
; FILE REFERENCE: ELITRA.011A
; CURRENT APPLICATION NUMBER: US/09/815,242
; CURRENT FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; NUMBER OF SEQ ID NOS: 14110
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 7863
; LENGTH: 2853
; TYPE: DNA
; ORGANISM: Pseudomonas aeruginosa
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)...(2853)
US-09-815-242-7863

Alignment Scores:
Pred. No.: 67.8 Length: 2853
Score: 66.00 Matches: 17
Percent Similarity: 46.94% Conservative: 6
Best Local Similarity: 34.69% Mismatches: 16
Query Match: 20.18% Indels: 10
DB: 10 Gaps: 3

09-833799-13B (1-57) x US-09-815-242-7863 (1-2853)

Qy 4 ProValLysGlyProValSerThrLysProGlySerCysProIleLeuLeuLeuArgCys 23
Db 313 CCCGCCGAGGTGCTGGCGGCTCACGCCCTGGG-----CGCCGA 275
Qy 24 AlaMetLeuAsnProProAsn-----ArgCysLeuLysAspThrAspCysProGlyIle 41
Db 274 GCTGGCGCTCGACCCACCATCTGGTGGCGATGCCGGCGGTGATCGG---TGCCCGGCTGCC 218
Qy 42 LysLysCysCysGluGlySerCysGly 50
Db 217 ACAGGGTGTTCGGGCCCTGCATGCGGC 191
RESULT 72
US-09-891-216-6/c
; Sequence 6, Application US/09891216

QY 3 GluProValLysGlyProValSerThrLysProGlySerCysProIleLe-----Leu 20
 ||||| ||||| ||||| ||||| :::
DB 6056 GAGTGTGCGCCCTTCCCGGCTCTTGCTCTCCTGGTACCTGTCAGAATTGGAGGGATCC 6115

QY 21 IleArgCysAlaMetLeuAsnProProAsnArgCysLeuLysAspThrAspCysProGly 40
Db 6116 TTCAGATGC-----ATCTGTCCCCCAGGGTATGAAGTAAAGCGAGAACTGCATTGAT 6169
QY 41 IleLysLysCysCysGlu-----GlySerCys----- 49
Db 6170 ATAAATGAATGTGATGAAGATCCCAACATTGTCTTTTGGTTCTCTGTAATAACTCCA 6229
QY 50 ---GlyMetAlaCysPheValPro 56
Db 6230 GGGGGCTTCCAGTGCCTCTGCCCC 6253

RESULT 75

US-10-044-090-47
; Sequence 47, Application US/10044090
; Patent No. US20020137081A1
; GENERAL INFORMATION:
; APPLICANT: Olga Bandman
; TITLE OF INVENTION: GENES DIFFERENTIALLY EXPRESSED IN VASCULAR TISSUE ACTIVATION
; FILE REFERENCE: PA-0028 US
; CURRENT APPLICATION NUMBER: US/10/044,090
; CURRENT FILING DATE: 2002-01-09
; NUMBER OF SEQ ID NOS: 850
; SOFTWARE: PERL Program
; SEQ ID NO 47
; LENGTH: 10284
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; OTHER INFORMATION: Incyte ID No. US20020137081A1 1098479.1
US-10-044-090-47

Alignment Scores:

Pred. No.:	370	Length:	10284
Score:	66.00	Matches:	22
Percent Similarity:	41.18%	Conservative:	6
Best Local Similarity:	32.35%	Mismatches:	24
Query Match:	20.18%	Indels:	16
DB:	12	Gaps:	4

09-833799-13B (1-57) x US-10-044-090-47 (1-10284)

QY 3 GluProValLysGlyProValSerThrLysProGlySerCysProIleIle-----Leu 20
Db 6056 GAGTGTGTGCGCCCTCCCGGCTCTTGGCTCTCCTGTACCTGTGAGAAATTTGAGGGATCC 6115
QY 21 IleArgCysAlaMetLeuAsnProProAsnArgCysLeuLysAspThrAspCysProGly 40
Db 6116 TTCAGATGC-----ATCTGTCCCCCAGGGTATGAAGTAAAGCGAGAACTGCATTGAT 6169
QY 41 IleLysLysCysCysGlu-----GlySerCys----- 49
Db 6170 ATAAATGAATGTGATGAAGATCCCAACATTGTCTTTTGGTTCTCTGTAATAACTCCA 6229
QY 50 ---GlyMetAlaCysPheValPro 56
Db 6230 GGGGGCTTCCAGTGCCTCTGCCCC 6253

Search completed: February 15, 2003, 22:01:09
Job time : 181 secs

GenCore version 5.1.4 p5 4578
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 15, 2003, 21:14:47 ; Search time 704.5 Seconds
(without alignments)
8509.834 Million cell updates/sec

Title: 09-833799-13A
Perfect score: 206
Sequence: 1 AATTCGAGCTCGGTACCATA.....GTTTCGTTCCACAATAATAG 206

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 2054640 seqs, 14551402878 residues
Total number of hits satisfying chosen parameters: 4109280

Minimum DB seq length: 0
Maximum DB seq length: 20000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 100 summaries

Database : GenEmbl:
1: gb_ba:
2: gb_htg:
3: gb_in:
4: gb_om:
5: gb_ov:
6: gb_pat:
7: gb_ph:
8: gb_pl:
9: gb_pr:
10: gb_ro:
11: gb_sts:
12: gb_sy:
13: gb_un:
14: gb_vi:
15: em_ba:
16: em_fun:
17: em_hum:
18: em_in:
19: em_mu:
20: em_om:
21: em_or:
22: em_ov:
23: em_pat:
24: em_ph:
25: em_pl:
26: em_ro:
27: em_sts:
28: em_un:
29: em_vi:
30: em_htg_hum:
31: em_htg_inv:
32: em_htg_other:
33: em_htg_mus:
34: em_htg_pln:
35: em_htg_rod:
36: em_htg_mam:
37: em_htg_vrt:
38: em_sy:
39: em_htgo_hum:
40: em_htgo_mus:
41: em_htgo_other:

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB	ID	Description
1	206	100.0	206	6	A31086	A31086 DNA fragmen
2	202	98.1	206	6	A31087	A31087 DNA fragmen
3	127	61.7	8598	6	E16036	E16036 CDNA encodi
4	125.4	60.9	177	6	I96047	I96047 Sequence 3
5	125.4	60.9	177	6	I96049	I96049 Sequence 7
6	125.2	60.8	480	6	AX397178	AX397178 Sequence
7	124.6	60.5	434	6	AX041087	AX041087 Sequence
8	124.6	60.5	478	9	HSANTLEUP	Z18538 H.sapiens e
9	124.6	60.5	504	6	A31090	A31090 DNA fragmen
10	124.6	60.5	571	6	I74316	I74316 Sequence 14
11	124.6	60.5	582	9	BC010952	BC010952 Homo sapi
12	123.8	60.1	177	6	I96048	I96048 Sequence 5
13	123.2	59.8	1878	9	HUMELAFIN	D13156 Human gene
14	123.2	59.8	2309	6	AX333644	AX333644 Sequence
15	123.2	59.8	2309	6	AX334316	AX334316 Sequence
16	123.2	59.8	2309	9	HUMPREELAS	L10343 Homo sapien
17	123.2	59.8	2309	9	S58717	S58717 pre-elafin=
18	123.2	59.8	99747	9	HSJ172H20	AL049767 Human DNA
19	123	59.7	347	6	A31089	A31089 DNA fragmen
20	121	58.7	321	6	I74313	I74313 Sequence 11
21	120.8	58.6	171	6	A31088	A31088 DNA fragmen
22	84.6	41.1	573	4	BOT223216	AJ223216 Bos tauru
23	75.2	36.5	3693	4	BTGWAPE	D50319 Pig DNA for
24	7					A31082 oligonucleo
25	7					AB003281 Phacochoe
26	73.1					D50322 Pig mRNA fo
27	73.1					D83668 Sus scrofa
28	71.1					AJ223215 Macaca mu
29	69					A31081 oligonucleo
30	68					A31080 oligonucleo
31	67.8					AB003283 Pecari ta
32	67.6					AB003284 Hippopota
33	67					A31083 oligonucleo
34	66.2					D17755 Porcine DNA
35	66.2					E06782 cDNA contai
36	66.2					D83667 Sus scrofa
37	66.2					D50320 Pig DNA for
38	66					A31079 oligonucleo
39	64					A31078 oligonucleo
40	63					AB003282 Phacochoe
41	60.4					D50323 Pig mRNA fo
42	54.6					AB011010 Bos tauru
43	43.4					D50321 Pig DNA for
44	42.6					AB003285 Sus scrof
45	39					AR067991 Sequence
46	38.4					AB042257 Cavia por
47	35.6					AC125967 Rattus no
48	35.4					AC006054 Homo sapi
49	35.2					AC096647 Homo sapi
50	34.8					M57446 Porcine ant
51	34.6					U88093 Mus musculu
52	34.6					U94341 Mus musculu
53	34.6					BC028509 Mus muscu
54	34.6					U73004 Mus musculu
55	34.2					AC110686 Rattus no
56	34					AC100679 Mus muscu
57	34					AL671891 Mus muscu
58	33.8					AL136380 Human DNA
59	33.8					AC116110 Mus muscu
60	33.6					AF002719 Mus muscu
61	33.6					AL662897 Human DNA
62	33.6					AL590429 Mouse DNA
63	33.6					AC021992 Homo sapi
64	33.4					AL035495 Human DNA
65	33.4					AC104349 Homo sapi

```
C 66 33.4 16.2 232793 2 AC107719 Mus muscu
67 33.2 16.1 4090 10 AF205374 Mus muscu
68 33.2 16.1 236121 2 AC126796 Mus muscu
69 33 16.0 17515 10 AF375596S2 Mus muscu
70 33 16.0 40223 9 AC119403 Homo sapi
71 33 16.0 42665 9 AC011559 Homo sapi
72 33 16.0 213703 2 AC119827 Homo sapi
73 32.8 15.9 143620 9 AC117109 Rattus no
74 32.6 15.8 120714 2 AC123297 Rattus no
75 32.6 15.8 142597 2 AC117109 Rattus no
76 32.6 15.8 176463 9 AC006377 Homo sapi
77 32.4 15.7 29188 9 HS77019 Human DNA
78 32.4 15.7 184007 9 AC116099 Homo sapi
79 32.4 15.7 188374 2 AC124042 Oryctolag
80 32.4 15.7 189718 9 AC012181 Homo sapi
81 32.2 15.6 324 6 I08404 Sequence 1
82 32.2 15.6 89811 2 AC108282 Rattus no
83 32.2 15.6 96807 2 AF205590 Homo sapi
84 32.2 15.6 112184 9 AC004054 Homo sapi
85 32.2 15.6 146469 2 AL355991 Homo sapi
86 32.2 15.6 172898 2 AL844551 Mus muscu
87 32.2 15.6 219622 9 AC009630 Homo sapi
88 32.2 15.6 273515 2 AC121844 Mus muscu
89 32 15.5 291 4 BOT223217 Bos tauru
90 32 15.5 115142 2 OSIG00038 Oryza sat
91 32 15.5 134884 2 AC016600 Homo sapi
92 32 15.5 138868 2 AC022375 Homo sapi
93 32 15.5 166976 2 AC034274 Homo sapi
94 32 15.5 170944 2 AC027230 Homo sapi
95 32 15.5 172507 9 AC092960 Homo sapi
96 32 15.5 230788 2 AC095410 Rattus no
97 31.8 15.4 2439 8 AF317473 Candida a
98 31.8 15.4 133213 2 AC109059 Rattus no
99 31.8 15.4 153553 9 AC008627 Homo sapi
100 31.8 15.4 158832 2 AC097950 Rattus no
```

ALIGNMENTS

```
RESULT 1
A31086 A31086 206 bp DNA linear PAT 21-AUG-1995
LOCUS DNA fragment from patent EP0402068.
DEFINITION A31086
ACCESSION A31086
VERSION A31086.1 GI:1249300
KEYWORDS synthetic construct.
SOURCE synthetic construct.
ORGANISM artificial sequences.
REFERENCE 1 (bases 1 to 206)
AUTHORS Christophers,E., Schroder,J.M., Pioli,D., Wiedow,O. and Edge,M.D.
TITLE Polypeptides and polypeptide analogues with inhibitory activity
against human elastase
JOURNAL Patent: EP 0402068-A 27 12-DEC-1990;
IMPERIAL CHEMICAL INDUSTRIES PLC
FEATURES
source
1 .206
/organism="synthetic construct"
/db_xref="taxon:32630"
BASE COUNT 51 a 46 c 44 g 65 t
ORIGIN
|||||
```

```
Query Match 100.0%; Score 206; DB 6; Length 206;
Best Local Similarity 100.0%; Pred. No. 2.5e-53;
Matches 206; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 AATTCGAGCTCGGTACCATACCTGCATATGCTCAAGAACCAAGTTAAAGTCTGTGCTA 60
|||||
Db 1 AATTCGAGCTCGGTACCATACCTGCATATGCTCAAGAACCAAGTTAAAGTCTGTGCTA 60
|||||
QY 61 CTAAGCCAGGTTCTTGTCTTATTATCTTGATTCGTTGGCTATGTTAAACCCACCTAAC 120
|||||
```

```
Db 61 CTAAGCCAGGTTCTTGTCTTATTATCTTGATTCGTTGGCTATGTTAAACCCACCTAAC 120
QY 121 GTTGTGTTGAAGGACACTGATTGTCCAGGTATCAAAAAGTCTGTGAAGTTCTCGGTA 180
|||||
Db 121 GTTGTGTTGAAGGACACTGATTGTCCAGGTATCAAAAAGTCTGTGAAGTTCTCGGTA 180
|||||
QY 181 TGGCTTGTTCGTTCCACAAATAATAG 206
|||||
Db 181 TGGCTTGTTCGTTCCACAAATAATAG 206
|||||
```

```
RESULT 2
A31087/c A31087 206 bp DNA linear PAT 21-AUG-1995
LOCUS DNA fragment from patent EP0402068.
DEFINITION A31087
ACCESSION A31087
VERSION A31087.1 GI:1247248
KEYWORDS synthetic construct.
SOURCE synthetic construct.
ORGANISM artificial sequences.
REFERENCE 1 (bases 1 to 206)
AUTHORS Christophers,E., Schroder,J.M., Pioli,D., Wiedow,O. and Edge,M.D.
TITLE Polypeptides and polypeptide analogues with inhibitory activity
against human elastase
JOURNAL Patent: EP 0402068-A 28 12-DEC-1990;
IMPERIAL CHEMICAL INDUSTRIES PLC
FEATURES
source
1 .206
/organism="synthetic construct"
/db_xref="taxon:32630"
BASE COUNT 64 a 45 c 47 g 50 t
ORIGIN
|||||
```

Query Match 98.1%; Score 202; DB 6; Length 206;
Best Local Similarity 100.0%; Pred. No. 4.4e-52;
Matches 202; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
QY 5 CGAGCTCGGTACCATACCTGCATATGCTCAAGAACCAAGTTAAAGTCTGTGCTACTAA 64
|||||
Db 206 CGAGCTCGGTACCATACCTGCATATGCTCAAGAACCAAGTTAAAGTCTGTGCTACTAA 147
|||||
QY 65 GCCAGGTTCTTGTCTTATTATCTTGATTCGTTGGCTATGTTAAACCCACCTAACCGTTG 124
|||||
Db 146 GCCAGGTTCTTGTCTTATTATCTTGATTCGTTGGCTATGTTAAACCCACCTAACCGTTG 87
|||||
QY 125 TTTGAAGGACACTGATTGTCCAGGTATCAAAAAGTCTGTGAAGTTCTCGGTTATGGC 184
|||||
Db 86 TTTGAAGGACACTGATTGTCCAGGTATCAAAAAGTCTGTGAAGTTCTCGGTTATGGC 27
|||||
QY 185 TTGTTTCGTTCCACAAATAATAG 206
|||||
Db 26 TTGTTTCGTTCCACAAATAATAG 5
|||||
```

```
RESULT 3
E16036 E16036 8598 bp DNA linear PAT 28-JUL-1999
LOCUS cDNA encoding improved Pichia elafin.
DEFINITION E16036
ACCESSION E16036
VERSION E16036.1 GI:5710719
KEYWORDS JP 1998127292-A/1.
SOURCE Pichia pastoris.
ORGANISM Pichia pastoris
```

```
REFERENCE Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
AUTHORS Taniyama,M., Yamamoto,T., Okawa,N. and Zushi,M.
TITLE ELAFINS-EXPRESSION VECTOR AND PRODUCTION OF ELAFINS BY UTILIZING
THE SAME
JOURNAL Patent: JP 1998127292-A 1 19-MAY-1998;
TSUMURA & CO
COMMENT OS Pichia pastoris
```

PN JP 1998127292-A/1
PD 19-MAY-1998
PF 31-OCT-1996 JP 1996304233
PI TANIYAMA MITSUE, YAMAMOTO TAKASHI, OKAWA NORIYUKI, PI ZUSHI
MAKOTO
PC C12N15/09,C07H21/04,C07K14/39,C12N1/19,C12P21/02,(C12N1/19, PC
C12R1:84),
PC (C12P21/02,C12R1:84);
CC strandedness: Double;
CC topology: Linear;
CC hypothetical: No;
CC anti-sense: No;
FH Key Location/Qualifiers

FT source 1..8598
FT /organism='Pichia pastoris'
FT /clone='pPIC9/BLF25L'
FT CDS 949..1377
FT /product='improved Pichia elafin' FT
FT sig_peptide 949..1203
FT mat_peptide 1204..1374
FT /product='improved Pichia elafin' FT
FT misc_feature replace(1276..1278,'atg')
FT /note='native Pichia elafin'.
FT Location/Qualifiers

FEATURES
source
1..8598
/organism="Pichia pastoris"
/db_xref="taxon:4922"
BASE COUNT 2306 a 2025 c 1920 g 2347 t
ORIGIN

Query Match 61.7%; Score 127; DB 6; Length 8598;
Best Local Similarity 82.9%; Pred. No. 9.1e-29;
Matches 145; Conservative 0; Mismatches 30; Indels 0; Gaps 0;

QY 30 GCTCAAGAACCCAGTTAAAGGTCCTGTGTCTACTAAGCCAGGTTCTTGTCCCTATTATCTTG 89
|||
Db 1204 GCTCAAGAACCCAGTTAAAGGTCCTGTGTCTACTAAGCCAGGTTCTTGTCCCTATTATCTTG 1263
|||
QY 90 ATTCGTTGCGCTATGTTAAACCCACCTAACCGTTGTTGAAGGACACTGATTGTCCAGGT 149
|||
Db 1264 ATCCGTCGCTTTGCTGAACCCGCGGAACCGTTGTCTGAAAGACACTGACTGCCCGGGT 1323
|||
QY 150 ATCAAAAAGTGTGTGAAGGTCCTGCGGTATGGCTTGTTCGTTCCACAATAAT 204
|||
Db 1324 ATCAAAAATGCTGCGAAGGTTCTTTCGCGGTATGGCATGCTTCGTTCCGCGAGTAGT 1378
|||

RESULT 4
I96047
LOCUS I96047 177 bp DNA linear PAT 01-DEC-1998
DEFINITION Sequence 3 from patent US 5734014.
ACCESSION I96047
VERSION I96047.1 GI:3940517
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 177)
AUTHORS Ishima,Y., Okawa,N., Yoshida,M., Amagaya,S. and Kaji,A.
TITLE Elafin derivative
JOURNAL Patent: US 5734014-A 3 31-MAR-1998;
FEATURES Location/Qualifiers
source 1..177
/organism="unknown"
BASE COUNT 39 a 49 c 48 g 41 t
ORIGIN

Query Match 60.9%; Score 125.4; DB 6; Length 177;
Best Local Similarity 82.3%; Pred. No. 3.1e-28;
Matches 144; Conservative 0; Mismatches 31; Indels 0; Gaps 0;

QY 30 GCTCAAGAACCCAGTTAAAGGTCCTGTGTCTACTAAGCCAGGTTCTTGTCTCTATTATCTTG 89

Db 1 GCACAGGAACCCAGTTAAAGGTCCTGCGGTGTCGACCAAAACCCGGCTCTTGCCCGATTATCCTG 60
|||
QY 90 ATTCGTTGCGCTATGTTAAACCCACCTAACCGTTGTTTGAAGGACACTGATTGTCCAGGT 149
|||
Db 61 ATCCGTCGCGCTTTGCTGAACCCGCGGAACCGTTGTCTGAAAGACACTGACTGCCCGGGT 120
|||
QY 150 ATCAAAAAGTGTGTGAAGGTTCTGCGGTATGGCTTGTTCGTTCCACAATAAT 204
|||
Db 121 ATCAAAAATGCTGCGAAGGTTCTTTCGCGGTATGGCATGCTTCGTTCCGCGAGTAGT 175
|||

RESULT 5
I96049
LOCUS I96049 177 bp DNA linear PAT 01-DEC-1998
DEFINITION Sequence 7 from patent US 5734014.
ACCESSION I96049
VERSION I96049.1 GI:3940519
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 177)
AUTHORS Ishima,Y., Okawa,N., Yoshida,M., Amagaya,S. and Kaji,A.
TITLE Elafin derivative
JOURNAL Patent: US 5734014-A 7 31-MAR-1998;
FEATURES Location/Qualifiers
source 1..177
/organism="unknown"
BASE COUNT 40 a 49 c 47 g 41 t
ORIGIN

Query Match 60.9%; Score 125.4; DB 6; Length 177;
Best Local Similarity 82.3%; Pred. No. 3.1e-28;
Matches 144; Conservative 0; Mismatches 31; Indels 0; Gaps 0;

QY 30 GCTCAAGAACCCAGTTAAAGGTCCTGTGTCTACTAAGCCAGGTTCTTGTCCCTATTATCTTG 89
|||
Db 1 GCACAGGAACCCAGTTAAAGGTCCTGCGGTGTCGACCAAAACCCGGCTCTTGCCCGATTATCCTG 60
|||
QY 90 ATTCGTTGCGCTATGTTAAACCCACCTAACCGTTGTTGAAGGACACTGATTGTCCAGGT 149
|||
Db 61 ATCCGTCGCGCTATTCTGAACCCGCGGAACCGTTGTCTGAAAGACACTGACTGCCCGGGT 120
|||
QY 150 ATCAAAAAGTGTGTGAAGGTCCTGCGGTATGGCTTGTTCGTTCCACAATAAT 204
|||
Db 121 ATCAAAAATGCTGCGAAGGTTCTTTCGCGGTATGGCATGCTTCGTTCCGCGAGTAGT 175
|||

RESULT 6
AX397178/c
LOCUS AX397178 480 bp DNA linear PAT 18-MAY-2002
DEFINITION Sequence 1393 from Patent WO0212328.
ACCESSION AX397178
VERSION AX397178.1 GI:21067925
KEYWORDS
SOURCE human.
ORGANISM Homo sapiens
REFERENCE 1
AUTHORS King,G.E., Meagher,M.J., Xu,J. and Secrist,H.
TITLE Compositions and methods for the therapy and diagnosis of colon cancer
JOURNAL Patent: WO 0212328-A 1393 14-FEB-2002;
FEATURES CORIXA CORPORATION (US)
Location/Qualifiers
source 1..480
/organism="Homo sapiens"
/db_xref="taxon:9606"
BASE COUNT 113 a 105 c 139 g 120 t 3 others
ORIGIN

REFERENCE	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
AUTHORS	1 (bases 1 to 582)
TITLE	Strausberg, R.
JOURNAL	Direct Submission
	Submitted (23-JUL-2001) National Institutes of Health, Mammalian
	Gene Collection (MGC), Cancer Genomics Office, National Cancer
	Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
	USA
REMARK	NIH-MGC Project URL: http://mgc.nci.nih.gov
COMMENT	Contact: MGC help desk
	Email: cgapbs-r@mail.nih.gov
	Tissue Procurement: ATCC
	cdNA Library Preparation: CLONTECH Laboratories, Inc.
	cdNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
	DNA Sequencing by: Institute for Systems Biology
	http://www.systemsbiology.org
	contact: amadan@systemsbiology.org
	Anup Madan, Rachel Dickhoff, Jessica Fahey, Stephanie Ford, Juli
	Greene, Mark Ketterman and Anuradha Madan

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>
Series: IRAL Plate: 19 Row: d Column: 7
This clone was selected for full length sequencing because it passed the following selection criteria: matched mRNA gi: 4505786.

FEATURES
source

```
1. 582 /organism="Homo sapiens"  
/db_xref="taxon:9606"  
/clone="MGC:13613 IMAGE:"  
/tissue_type="Brain, gli  
/clone_lib="NIH MGC_57"  
/lab_host="DH10B"  
/note="Vector: pDNR-LIB"  
26. 379
```

CDs

```

/codon_start=1
/product="Similar to protease inhibitor 3, skin-derived
(SKALP)"
/protein_id="AAH10952.1"
/db_xref="GI:15012095"
/translation="MRASSEFLIVVFLIAGTLVLEAAVTGVPVKQDTVKGRVPFNGQ
DPVKGQSVKQDKVKAQEPVKGPVSTKPGSCPIILIRCAMLPNPNRCLKDTCPGIK
KCCGEGSCGMACEVRQ"

```

BASE COUNT	148 a	166 c	134 g	134 t
ORIGIN				

```
Query Match      60.5%; Score 124.6; DB 9; Length 582;
Best Local Similarity 81.0%; Pred. No. 5.3e-28;
Matches 145: Conservative 0; Mismatches 34; Indels 0
```

[illegible]

RESULT 12	
I96048	
LOCUS	
DEFINITION	
ACCESSION	
VERSION	
KEYWORDS	
SOURCE	
ORGANISM	

I96048 177 bp DNA linear PAT 01-DEC-1998
Sequence 5 from patent US 5734014.
I96048
I96048.1 GI:3940518

Unknown.
Unknown.
Unclassified

from Fig. 4"
/codon_start=1
/product="pre-elafin"
/protein_id="AAB26371.1"
/db_xref="GI:299841"
/translation="MRASSFLIVVFLIAGTLVLEAAVTGVVKQDTPVGRVPFNGQ
DPVKGVSVKGDVKVKAQEPVKGPVSTKPGSCPIILIRCAMLPNRRCLKDITDCPIK
KCEGSGCMACFPVQ"
BASE COUNT 595 a 537 c 603 g 572 t 2 others
ORIGIN

Query Match 59.8%; Score 123.2; DB 9; Length 2309;
Best Local Similarity 81.2%; Pred. No. 1.4e-27;
Matches 143; Conservative 0; Mismatches 33; Indels 0; Gaps 0;

QY 28 ATGCTCAAGAACCCAGTTAAAGGTCCTGTGTCTACTAAGCCAGGTTCTTGTCCTATTATCT 87
Db 1552 AAGCGCAAGAGCCAGTCAAAAGGTCAGTCTCCACTAAGCCTGGCTCCTGCCCATTTATCT 1611
QY 88 TGATTCGTTGCGCTATGTTAAACCCACCTAACCGTTGTTGAAGGACACTGATTGTCCAG 147
Db 1612 TGATCCGTGCGCCATGTTGAATCCCCCTAACCGTCTGTTGAAAGATACACTGACTGCCAG 1671
QY 148 GTATCAAAAAGTGTGTGAAGGTTCTCTGCGGTATGCTTGTTCGTTCCACAATAA 203
Db 1672 GAATCAAGAAGTGTGTGAAGGCTCTTGGCGGATGGCCTGTTTCGTTCCCCAGTGA 1727

RESULT 18
HSJ172H20
LOCUS
DEFINITION
Human DNA sequence from clone RP1-172H20 on chromosome 20q12-13.12
Contains SEMG1 and SEMG2 genes for senenogelin I and II, the PI3
gene for protease inhibitor 3, skin-derived (SKALP), ESTs, STSs and
GSSs, complete sequence.
AL049767
AL049767.12 GI:11345541
HTG; PI3; SEMG1; SEMG2; SKALP.
human.
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 99747)
Sehra,H.
Direct Submission
Submitted (29-MAR-2001) Sanger Centre, Hinxton, Cambridgeshire,
CB10 1SA, UK. E-mail enquiries: humquery@sanger.ac.uk
requests: clonerequest@sanger.ac.uk
On Nov 27, 2000 this sequence version replaced gi:10198636.
During sequence assembly data is compared from overlapping clones.
Where differences are found these are annotated as variations
together with a note of the overlapping clone name. Note that the
variation annotation may not be found in the sequence submission
corresponding to the overlapping clone, as we submit sequences with
only a small overlap as described above.
The following abbreviations are used to associate primary accession
numbers given in the feature table with their source databases:
Em., EMBL; Sw., SWISSPROT; Tr., TREMBL; Wp., WORMPEP; Information
on the WORMPEP database can be found at
http://www.sanger.ac.uk/Projects/C_elegans/wormpep This sequence
was generated from part of bacterial clone contigs of human
chromosome 20, constructed by the Sanger Centre Chromosome 20
Mapping Group. Further information can be found at
http://www.sanger.ac.uk/HGP/Chr20
IMPORTANT: This sequence is not the entire insert of clone
RP1-172H20 It may be shorter because we sequence overlapping
sections only once, except for a 100 base overlap.
The true right end of clone RP1-172H20 is at 99747 in this
sequence. The true left end of clone RP1-30012 is at 28594 in this
sequence. The true right end of clone RP1-211D12 is at 100 in this
sequence. This sequence was finished as follows unless otherwise
noted: all regions were either double-stranded or sequenced with an
alternate chemistry or covered by high quality data (i.e., phred

quality >= 30); an attempt was made to resolve all sequencing
problems, such as compressions and repeats; all regions were
covered by at least one plasmid subclone or more than one M13
subclone; and the assembly was confirmed by restriction digest.
RP1-172H20 is from the library RPCI-1 constructed by the group of
Pieter de Jong. For further details see
http://www.chori.org/bacpac/home.htm
VECTOR: pCYPAC2.

FEATURES
Location/Qualifiers
source
1..99747
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="20"
/map="q12-q13.12"
/clone="RP1-172H20"
/clone_lib="RPCI-1"
complement(1..100)
/note="match: STS: Em:HS211D12S"
1116..1993
/note="L1MB1 repeat: matches 5294. .6159 of consensus"
2001..2042
/note="21 copies 2 mer ta 83% conserved"
2272..2376
/note="L1 repeat: matches 5193. .5300 of consensus"
2377..2408
/note="16 copies 2 mer ta 84% conserved"
2411..2446
/note="18 copies 2 mer at 80% conserved"
complement(2469..2951)
/note="match: GSS: Em:AQ809877"
2581..2789
/note="MER20 repeat: matches 7. .218 cf consensus"
2982..3563
/note="match: GSS: Em:AQ386436"
2989..3136
/note="L1MB4 repeat: matches 3850. .3998 of consensus"
3780..4090
/note="AluX repeat: matches 1. .312 cf consensus"
4294..4592
/note="AluX repeat: matches 1. .309 of consensus"
4965..5002
/note="19 copies 2 mer gt 94% conserved"
5291..5487
/note="HAL1 repeat: matches 1501. .1692 of consensus"
5569..6048
/note="L1MB1 repeat: matches 5696. .6172 of consensus"
6049..6158
/note="MER20 repeat: matches 109. .218 of consensus"
6160..6228
/note="L2 repeat: matches 2021. .2087 of consensus"
6286..6466
/note="L2 repeat: matches 2329. .2497 of consensus"
6548..6731
/note="L1MB4 repeat: matches 7650. .7849 of consensus"
6735..8399
/note="L1MB8 repeat: matches 4484. .6153 of consensus"
8460..8581
/note="FLAM_A repeat: matches 1. .122 of consensus"
8589..10634
/note="L1MB3e repeat: matches 22. .1826 of consensus"
10656..10979
/note="MER74A repeat: matches 248. .557 of consensus"
11205..11377
/note="AluSg/x repeat: matches 136. .308 of consensus"
11408..11859
/note="LTR16B repeat: matches 2. .461 of consensus"
11973..12160
/note="Charlie2 repeat: matches 3534. .3727 of consensus"
14238..14548
/note="match: STS: Em:G08051"
14337..14396
/note="30 copies 2 mer ct 75% conserved"
14492..17136

repeat_region /note="HERVL repeat: matches 4. .2665 of consensus" 17140. .17589

repeat_region /note="MLT2B repeat: matches 1. .444 of consensus" 17590. .17899

repeat_region /note="L1 repeat: matches 4299. .4610 of consensus" 17898. .18682

repeat_region /note="L1M1 repeat: matches -1390. -.578 of consensus" 19077. .19213

repeat_region /note="L1ME repeat: matches 5694. .5824 of consensus" 20761. .20831

repeat_region /note="L1M4b repeat: matches -259. -.184 of consensus" 20853. .21014

repeat_region /note="L1MA9 repeat: matches 6102. .6270 of consensus" 21044. .21573

repeat_region /note="L1M4b repeat: matches -263. .330 of consensus" 21634. .21829

repeat_region /note="L1M4b repeat: matches 330. .524 of consensus" 22891. .23199

repeat_region /note="AluX repeat: matches 1. .309 of consensus" 25460. .25629

repeat_region /note="L1M3a repeat: matches 202. .531 of consensus" 28959. .29336

repeat_region /note="L1M1 repeat: matches 469. .859 of consensus" 29337. .29711

repeat_region /note="THE1C repeat: matches 1. .371 of consensus" 29712. .31317

repeat_region /note="THE1C-internal repeat: matches 1. .1580 of consensus" 31318. .31729

repeat_region /note="THE1C repeat: matches 1. .371 of consensus" 31730. .31886

repeat_region /note="L1M1 repeat: matches 859. .987 of consensus" 33209. .34261

repeat_region /note="L1M3d repeat: matches 859. .1833 of consensus" 34232. .34942

repeat_region /note="L1M1 repeat: matches 1560. .2657 of consensus" 34941. .36117

repeat_region /note="L1 repeat: matches 3935. .5142 of consensus" 36118. .36412

repeat_region /note="AluJo repeat: matches 9. .292 of consensus" 36430. .37302

repeat_region /note="L1 repeat: matches 2905. .3780 of consensus" 37299. .37425

repeat_region /note="L1 repeat: matches 4912. .5038 of consensus" 37434. .37469

repeat_region /note="L1P repeat: matches 2850. .2885 of consensus" 38377. .39183

repeat_region /note="L1M4c repeat: matches 1312. .1666 of consensus" 39261. .42088

repeat_region /note="L1PA15 repeat: matches 671. .3485 of consensus" 42057. .43098

repeat_region /note="L1PA5 repeat: matches 5104. .6143 of consensus" 43119. .45775

repeat_region /note="L1PA15 repeat: matches 3467. .6157 of consensus" 45840. .45956

repeat_region /note="L1ME repeat: matches 5530. .5647 of consensus" 46081. .46114

repeat_region /note="17 copies 2 mer tc 82% conserved" 46492. .46761

repeat_region /note="L1ME repeat: matches 5521. .5793 of consensus" 46828. .46915

repeat_region /note="L1ME3A repeat: matches 5300. .5388 of consensus" 46916. .47210

repeat_region /note="AluSq repeat: matches 1. .305 of consensus" 47211. .47769

misc_feature /note="L1ME3A repeat: matches 5388. .5937 of consensus" complement(48203. .48496)

repeat_region /note="match: GSS: Em:AQ275745" 48638. .48679

gene /note="21 copies 2 mer ac 97% conserved" 49535. .51104

mRNA /gene="PI3" join(49535. .49615,50475. .50750,50984. .51104) /gene="PI3" /product="dJ172H20.3 (protease inhibitor 3, skin-derived (SKALP))" /note="match: cDNAs: Em:Z18538 Em:AJ223218 Em:AJ223216 match: ESTs: Em:AA527557 Em:AA583567 Em:AA582866" /evidence=not_experimental join(49537. .49615,50475. .50749) CDS

Query Match 59.8%; Score 123.2; DB 9; Length 99747; Best Local Similarity 81.2%; Pred. No. 1.3e-27; Matches 143; Conservative 0; Mismatches 33; Indels 0; Gaps 0;

Qy 28 ATGCTCAAGAACCCAGTTAAAGGTCCTGTGTCTACTAAGCCAGGTTCTTGTCTCTATTATCT 87

Db 50574 AAGCGCAAGAGCCAGTCAAAGGTCAGTCTCCACTAAGCCTGGCTCCTGCCCATATCT 50633

Qy 88 TGATTCTGTCGGCTATGTTAAACCCACCTAACCGTTGTTTGAAGGACACTGATTGTCCAG 147

Db 50634 TGATCCGGTGGCCATGTTGAATCCCCCTAACCGCTGCTTGAAGATATACTGACTGCCAG 50693

Qy 148 GTATCAAAAAGTGCTGTGAAGTTCCTCGGTATGGCTTGTTCGTTCCACAATAA 203

Db 50694 GAATCAAGAAGTGCTGTGAAGGCTCTTGGGGATGGCCTGTTTCGTTCCCCAGTGA 50749

RESULT 19

A31089

LOCUS A31089 347 bp DNA linear PAT 21-AUG-1995

DEFINITION DNA fragment from patent EP0402068.

ACCESSION A31089

VERSION A31089.1 GI:1249302

KEYWORDS

SOURCE synthetic construct.

ORGANISM synthetic construct

REFERENCE 1 (bases 1 to 347)

AUTHORS Christophers,E., Schroder,J.M., Pioli,D., Wiedow,O. and Edge,M.D.

TITLE Polypeptides and polypeptide analogues with inhibitory activity against human elastase

JOURNAL Patent: EP 0402068-A 30 12-DEC-1990; IMPERIAL CHEMICAL INDUSTRIES PLC

FEATURES Location/Qualifiers

source 1. .347

BASE COUNT 65 a 118 c 78 g 84 t 2 others

ORIGIN

Query Match 59.7%; Score 123; DB 6; Length 347; Best Local Similarity 80.2%; Pred. No. 1.7e-27; Matches 142; Conservative 2; Mismatches 33; Indels 0; Gaps 0;

Qy 30 GCTCAAGAACCCAGTTAAAGGTCCTGTGTCTACTAAGCCAGGTTCTTGTCTCTATTATCTG 89

Db 1 GCGCAAGAGCCAGTCAAAGGTCAGTCTCCACTAAGCCTGGCTCCTGCCCATATCTTG 60

Qy 90 ATTCGTTGCGCTATGTTAAACCCACCTAACCGTTGTTTGAAGGACACTGATTGTCCAGGT 149

Db 61 ATCCGGTGGCCATGTTGAATCCCCCTAACCGCTGCTTGAAGATATACTGACTGCCAGGA 120

Qy 150 ATCAAAAAGTGTGAAGGTTCCCTCGGTATGGCTTGTTCGTTCCACAATAATAG 206

Db 121 ATHAAGAARTGCTGTGAAGGCTCTTGGGGATGGCCTGTTTCGTTCCCCAGTAGGAG 177

RESULT 20

I74313

LOCUS I74313 321 bp DNA linear PAT 03-APR-1998

DEFINITION Sequence 11 from patent US 5688641.

QY 33 CAAGAACACAGTTAAAGTCTGTGTCTACTAAGCCAGGTTCTTGTCTCTATTATCTTGATT 92
|||||
Db 235 CAAGGCCGAATCGGAGTCCACTCCTCACTAAGCTGGTCTCTGCCCGAGGTTCTGATC 294
QY 93 CGTTGCGCTATGTTAAACCCACCTAAACCGTTGTTGAAGGACACTGATTGTCCAGGTATC 152
Db 295 CGGTGTCCATGATGAACCCCTTAACCGATGTCTGAGGGATGCTCAGTCCCGAGGGTC 354
QY 153 AAAAAAGTGTGTGAAGTTCCTGCGGTATGCGTCTGTTTCGTTCCACAATAA 203
Db 355 AAGAAGTGTGTGAAGGCTCTTGTGGGAAGACCTGTATGGATCCCCAGTGA 405
RESULT 23
PIGWAPA
LOCUS PIGWAPA 3693 bp DNA linear MAM 14-APR-2000
DEFINITION Pig DNA for elafin, complete cds.
ACCESSION D50319
VERSION D50319.1 GI:1228057
KEYWORDS elafin.
SOURCE Sus scrofa DNA, clone_lib:EMBL SP6/T7 clone:lambda WAP-1.
ORGANISM Sus scrofa
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
REFERENCE 1 (sites)
AUTHORS Tamechika, I., Itakura, M., Saruta, Y., Furukawa, M., Kato, A.,
Tachibana, S. and Hirose, S.
TITLE Accelerated evolution in inhibitor domains of porcine elafin family
members
JOURNAL J. Biol. Chem. 271 (12), 7012-7018 (1996)
MEDLINE 96215132
REFERENCE 2 (bases 1 to 3693)
AUTHORS Hirose, S.
JOURNAL Unpublished
REFERENCE 3 (bases 1 to 3693)
AUTHORS Hirose, S.
TITLE Direct Submission
JOURNAL Submitted (18-APR-1995) Shigehisa Hirose, Tokyo Institute of
Technology, Department of Biological Sciences; 4259 Nagatsuta-cho,
Midori-ku, Yokohama, Kanagawa 226-8501, Japan
(E-mail: shirose@bio.titech.ac.jp, Tel: 045-924-5726,
Fax: 045-5824)
FEATURES
source
1. .3693
/organism="Sus scrofa"
/db_xref="taxon:9823"
/clone_lib="lambda WAP-1"
/clone_lib="EMBL SP6/T7"
<1. .710
/number=1
<1. .628
join(629. .710,1541. .1962)
/codon_start=1
/product="elafin"
/protein_id="BAA08854.1"
/db_xref="GI:1228058"
/translation="MRSRSFLVVLVVFLLVGLVLAQAAGRIRPKGKTKKILALVKG
QGPVRGKDQVKGQGVKQDGLGKSDPVKAQLPDKGQDLGKGEDSVKQDPPFKAQLPD
KLQDPVKAQPAIKRLILLITKPGSCPRILIRCLMVNPPNRLCLSDAQCPGLKKCEGFCG
KACMDPK"
629. .691
1777. .1959
/product="elafin"
711. .1540
/number=1
1541. .1963
/number=2
join(1963,2445. .2601)
1964. .2444
/number=2
2141. .2402
/note="SINE (PRE-1)"
sig_peptide
mat_peptide
intron
exon
3'UTR
intron
repeat_unit

exon 2445. .2601
/number=3
polyA_signal 2580. .2585
polyA_site 2601
repeat_region 3606. .3669
/note="TG repeat"
BASE COUNT 798 a 953 c 1112 g 830 t
ORIGIN
Query Match 36.5%; Score 75.2; DB 4; Length 3693;
Best Local Similarity 70.1%; Pred. No. 1.2e-12;
Matches 101; Conservative 0; Mismatches 43; Indels 0; Gaps 0;
QY 60 ACTAAGCCAGGTTCTTGTCTCTATTATCTTGTATTCGTTGGCTATGTTAAACCCACCTAAC 119
|||||
Db 1819 ACCAAGCCTGGCTCCTGCCAGGATTCTGATCCGTTGCTTGATGGTCAATCCCCCTAAC 1878
QY 120 CGTTGTTTGAAGGACACTGATGTCCAGGTATCAAAAAGTCTGTGAAGGTTCTCGCGGT 179
|||||
Db 1879 AGGTGTTTGAAGTATGCTCAGTCCAGGGCTCAAGAAAGTCTGTGAAGGCTTTTGGGG 1938
QY 180 ATGGCTTGTTCGTTCCACAATAA 203
|||||
Db 1939 AAGCCTGTATGGATCCCAAGTGA 1962
RESULT 24
A31082
LOCUS A31082 74 bp DNA linear PAT 21-AUG-1995
DEFINITION oligonucleotide ELI5 from patent EP0402068.
ACCESSION A31082
VERSION A31082.1 GI:1249296
KEYWORDS synthetic construct.
SOURCE synthetic construct.
ORGANISM artificial sequences.
REFERENCE 1 (bases 1 to 74)
AUTHORS Christophers, E., Schroder, J.M., Pioli, D., Wiedow, O. and Edge, M.D.
TITLE Polypeptides and polypeptide analogues with inhibitory activity
against human elastase
JOURNAL Patent: EP 0402068-A 22 12-DEC-1990;
IMPERIAL CHEMICAL INDUSTRIES PLC
FEATURES
source
1. .74
/organism="synthetic construct"
/db_xref="taxon:32630"
BASE COUNT 19 a 14 c 18 g 23 t
ORIGIN
Query Match 35.9%; Score 74; DB 6; Length 74;
Best Local Similarity 100.0%; Pred. No. 3.1e-12;
Matches 74; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 133 AACTGATTGTCCAGGTATCAAAAAGTCTGTGAAGGTTCTCGCGGTATGGCTGTTTCG 192
|||||
Db 1 ACACGTATTGTCCAGGTATCAAAAAGTCTGTGAAGGTTCTCGCGGTATGGCTGTTTCG 60
QY 193 TTCCACAATAATAG 206
|||||
Db 61 TTCCACAATAATAG 74
RESULT 25
AB003281
LOCUS AB003281 1128 bp DNA linear MAM 26-MAY-1999
DEFINITION Phacochoerus aethiopicus gene for elafin (trappin-2), partial cds.
ACCESSION AB003281
VERSION AB003281.1 GI:4887637
KEYWORDS elafin (trappin-2).
SOURCE Phacochoerus aethiopicus DNA.
ORGANISM Phacochoerus aethiopicus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Phacochoerus.

Hirose,S.
Evolution of the trappin multigene family in the Suidae
J. Biochem. 124 (3), 491-502 (1998)
98391820
PUBMED
REFERENCE 3 (bases 1 to 1034)
AUTHORS Hirose,S.
TITLE Direct Submission
JOURNAL Submitted (23-FEB-1996) Shigehisa Hirose, Tokyo Institute of Technology, Department of Biological Sciences; 4259 Nagatsuta-cho, Midori-ku, Yokohama, Kanagawa 226-8501, Japan
(E-mail:shirose@bio.titech.ac.jp, Tel:045-924-5726, Fax:045-924-5824)
COMMENT On Apr 29, 1997 this sequence version replaced gi:1209380.
Sequence updated (21-Apr-1997).
FEATURES
 source
 1..1034
 /organism="Sus scrofa"
 /db_xref="taxon:9823"
 /clone_lib="PCR product"
 <1..208
 /number=1
 intron
 1..24
 /note="PCR primer"
 exon
 209..559
 /number=2
 <209..558
 /note="exon2 part"
 /codon_start=3
 /product="elafin homolog"
 /protein_id="BAA12038.1"
 /db_xref="GI:1209381"
 /translation="RPGKGTGKTLALVKGQGPVRKDOVKGQPVKGDIGKSDPV
 KAQLPDKGQDPVKAQPAIKRLILLTKGSCPRILIRCLMVNPPNRCLSDAQCPGVKKK
 CEGFCGKDCMDPK"
 mat_peptide
 373..555
 /product="elafin homolog"
 /function="elastase specific inhibitor"
 /note="Mature protein has a characteristic
 disulfide-linked structure called four disulfide core or
 WAP motif"
 560..>1034
 /number=2
 1010..>1034
 /note="PCR primer"
BASE COUNT 263 a 262 c 298 g 211 t
ORIGIN
Query Match 35.7%; Score 73.6; DB 4; Length 1034;
Best Local Similarity 69.4%; Pred. No. 3.9e-12;
Matches 100; Conservative 0; Mismatches 44; Indels 0; Gaps 0;
QY 60 ACTAAGCCAGGTTCTTGTCTTATCTTATCTTGTGCTCGCTATGTTAAACCCACCTAAC 119
Db 415 ACCAAGCCTGGCTCCTGCTCCAGGATTCTGATCGCTGCTGATGGTCAATCCCCCTAAC 474
QY 120 CGTTGTTGAAGGACACTGATTGTCCAGGTATCAAAAAGTGTGTGAAGGTTCTGCGGT 179
Db 475 AGGTGTTGATGATGCTCAGTGCCAGGGGTCAAGAGTGTGTGAAGGCTTTTGCGGG 534
QY 180 ATGGCTTGTTCGTTCCACAATAA 203
Db 535 AAGGACTGTATGATGCCCAAGTGA 558
RESULT 28
MAM223215 MAM223215 270 bp mRNA linear PRI 15-DEC-1998
LOCUS Macaca mulatta mRNA for putative sTrappin-2 protein, partial.
DEFINITION AJ223215
ACCESSION AJ223215
VERSION AJ223215.1 GI:2764783
KEYWORDS proteinase inhibitor; putative; sTrappin-2 gene; transglutaminase substrate.

SOURCE
ORGANISM rhesus monkey.
Macaca mulatta
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Euthera; Primates; Catarrhini; Cercopitheidae;
Cercopitheciniae; Macaca.
REFERENCE 1 (bases 1 to 270)
AUTHORS Zeeuwen,P.L.J.M.
TITLE Direct Submission
JOURNAL Submitted (06-JAN-1998) Zeeuwen P.L.J.M., Dermatology, University Hospital Nijmegen, PO BOX 9101, 6500 HB Nijmegen, THE NETHERLANDS
REFERENCE 2 (bases 1 to 270)
AUTHORS Zeeuwen,P.L., Hendriks,W., de Jong,W.W. and Schalkwijk,J.
TITLE Identification and sequence analysis of two new members of the SKALP/elafin and SPAI-2 gene family. Biochemical properties of the transglutaminase substrate motif and suggestions for a new nomenclature
J. Biol. Chem. 272 (33), 20471-20478 (1997)
MEDLINE 97400522
PUBMED 9252357
REFERENCE 3 (bases 1 to 270)
AUTHORS Tamechika,I., Itakura,M., Saruta,Y., Furukawa,M., Kato,A., Tachibana,S. and Hirose,S.
TITLE Accelerated evolution in inhibitor domains of porcine elafin family members
J. Biol. Chem. 271 (12), 7012-7018 (1996)
MEDLINE 96215132
PUBMED 8636131
REFERENCE 4 (bases 1 to 270)
AUTHORS Molhuizen,H.O., Alkemade,H.A., Zeeuwen,P.L., de Jongh,G.J., Wieringa,B. and Schalkwijk,J.
TITLE SKALP/elafin: an elastase inhibitor from cultured human keratinocytes. Purification, cDNA sequence, and evidence for transglutaminase cross-linking
J. Biol. Chem. 268 (16), 12028-12032 (1993)
MEDLINE 93280175
PUBMED 7685029
FEATURES
 source
 1..270
 /organism="Macaca mulatta"
 /db_xref="taxon:9544"
 /tissue_type="skin"
 1..270
 /gene="sTrappin-2"
 <1..>270
 /gene="sTrappin-2"
 /function="transglutaminase substrate, putative proteinase inhibitor"
 /codon_start=2
 /protein_id="CAA11183.1"
 /db_xref="GI:2764784"
 /db_xref="SPTREMBL:O46643"
 /translation="VVVFLIAGMLVVEAAVTGVPVKGQDTPVGRVFPFNGQDPVKGVQS
 VKQDRVKGRGPVKGPVSTKPGSCPILIRCAMLNPPNRCLUKDTD"
 sig_peptide
 <1..43
 /gene="sTrappin-2"
 mat_peptide
 41..>270
 /gene="sTrappin-2"
 /product="unnamed"
 /function="transglutaminase substrate, putative proteinase inhibitor"
BASE COUNT 64 a 66 c 74 g 66 t
ORIGIN
Query Match 34.9%; Score 71.8; DB 9; Length 270;
Best Local Similarity 76.5%; Pred. No. 1.5e-11;
Matches 88; Conservative 0; Mismatches 27; Indels 0; Gaps 0;
QY 28 ATGCTCAAGAACCCAGTTAAAGTCTGTGTCTACTAAGCCAGGTTCTTGTCTCTATCT 87
Db 156 AAGGCGAGGGCCAGTCAAGGTCCTCCACTAAGCCTGGCTCTGCCCAATATCT 215
QY 88 TGATTCGTTGCGCTATGTTAAACCCACCTAACCGTTGTTGAAGGACACTGATG 142

Db 216 TGATCCGGTGGCCCATGTTGAATCCCCCTAACCGCTGCTTGAAAGATACTGACTG 270

RESULT 29
A31081/c
LOCUS A31081 69 bp DNA linear PAT 21-AUG-1995
DEFINITION oligonucleotide ELI4 from patent EP0402068.
ACCESSION A31081
VERSION A31081.1 GI:1249295
KEYWORDS
SOURCE
ORGANISM
synthetic construct.
synthetic construct
artificial sequences.
REFERENCE
1 (bases 1 to 69)
AUTHORS Christoffers,E., Schroder,J.M., Pioli,D., Wiedow,O. and Edge,M.D.
TITLE Polypeptides and polypeptide analogues with inhibitory activity
JOURNAL
JOURNAL
IMPERIAL CHEMICAL INDUSTRIES PLC
Patent: EP 0402068-A 21 12-DEC-1990;
IMPERIAL CHEMICAL INDUSTRIES PLC
FEATURES
Location/Qualifiers
1..69
/organism="synthetic construct"
/db_xref="taxon:32630"
BASE COUNT 26 a 12 c 16 g 15 t
ORIGIN

Query Match 33.5%; Score 69; DB 6; Length 69;
Best Local Similarity 100.0%; Pred. No. 1.1e-10;
Matches 69; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 71 TTCTTGCTCTATTATCTTGATTCGTCGCTATGTTAAACCCACCTAACCGTTGTTGAA 130
Db 69 TTCTTGCTCTATTATCTTGATTCGTCGCTATGTTAAACCCACCTAACCGTTGTTGAA 10

QY 131 GGACACTGA 139
Db 9 GGACACTGA 1

RESULT 30
A31080
LOCUS A31080 68 bp DNA linear PAT 21-AUG-1995
DEFINITION oligonucleotide ELI3 from patent EP0402068.
ACCESSION A31080
VERSION A31080.1 GI:1249294
KEYWORDS
SOURCE
ORGANISM
synthetic construct.
synthetic construct
artificial sequences.
REFERENCE
1 (bases 1 to 68)
AUTHORS Christoffers,E., Schroder,J.M., Pioli,D., Wiedow,O. and Edge,M.D.
TITLE Polypeptides and polypeptide analogues with inhibitory activity
JOURNAL
JOURNAL
IMPERIAL CHEMICAL INDUSTRIES PLC
Patent: EP 0402068-A 20 12-DEC-1990;
IMPERIAL CHEMICAL INDUSTRIES PLC
FEATURES
Location/Qualifiers
1..68
/organism="synthetic construct"
/db_xref="taxon:32630"
BASE COUNT 13 a 16 c 14 g 25 t
ORIGIN

Query Match 33.0%; Score 68; DB 6; Length 68;
Best Local Similarity 100.0%; Pred. No. 2.3e-10;
Matches 68; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 65 GCCAGGTTCTTGTCTATTATCTTGATTCGTCGCTATGTTAAACCCACCTAACCGTTG 124
Db 1 GCCAGGTTCTTGTCTATTATCTTGATTCGTCGCTATGTTAAACCCACCTAACCGTTG 60

QY 125 TTTGAAGG 132
Db 61 TTTGAAGG 68

RESULT 31
AB003283
LOCUS AB003283 1169 bp DNA linear MAM 26-MAY-1999
DEFINITION Pecari tajacu gene for trappin, partial cds.
ACCESSION AB003283
VERSION AB003283.1 GI:4887641
KEYWORDS trappin.
SOURCE Pecari tajacu DNA.
ORGANISM Pecari tajacu
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Cetartiodactyla; Suina; Tayassuidae; Pecari.
REFERENCE
1 (sites)
AUTHORS Furutani,Y., Kato,A., Yasue,H., Alexander,L.J., Beattie,C.W. and Hirose,S.
TITLE Evolution of the trappin multigene family in the Suidae
JOURNAL J. Biochem. 124 (3), 491-502 (1998)
MEDLINE 98391820
REFERENCE
2 (bases 1 to 1169)
AUTHORS Hirose,S.
TITLE Direct Submission
JOURNAL Submitted (21-APR-1997) Shigehisa Hirose, Tokyo Institute of Technology, Department of Biological Sciences; 4259 Nagatsuta-cho, Midori-ku, Yokohama, Kanagawa 226-8501, Japan
(E-mail:shiroyose@bio.titech.ac.jp, Tel:045-924-5726, Fax:045-924-5824)
FEATURES
Location/Qualifiers
1..1169
/organism="Pecari tajacu"
/db_xref="taxon:9829"
CDS
<212..675
/codon_start=3
/product="trappin"
/protein_id="BAA77827.1"
/db_xref="GI:4887642"
/translation="RPKGQGTKKGHALIKQDPPVRGQGPVKQDPPVKAQLPDKGQDLV
KGQDPPVKQDPPVKAQLPDKGQDLVKQDPPVKQDPPVKQDPPVKAELAVRRLV
LPRRKPGFCPMIKIRCALFNPNNRCLTDAGCPGARKCCIGSCGKACLNPVR"
212..>675
exon
BASE COUNT 305 a 289 c 330 g 245 t
ORIGIN

Query Match 32.9%; Score 67.8; DB 4; Length 1169;
Best Local Similarity 68.9%; Pred. No. 2.5e-10;
Matches 93; Conservative 0; Mismatches 42; Indels 0; Gaps 0;

QY 62 TAAGCCAGGTTCTTGTCTATTATCTTGATTCGTCGCTATGTTAAACCCACCTAACCG 121
Db 531 TAAGCCAGGTTCTTGTCTATTATCTTGATTCGTCGCTATGTTAAACCCACCTAACCG 590

QY 122 TTGTTGAAGGACACTGATTCGTCGCTATCAAAAGTCTGTGAAGTTCTCGGTTAT 181
Db 591 GTGTTGACCGATGCTGGTGCCAGGGCCAGGAAGTCTGCATAGGCTCTTGGCGGAA 650

QY 182 GGCTTGTTCGTTCC 196
Db 651 GGCTTGTTCGTTCC 665

RESULT 32
AB003284
LOCUS AB003284 734 bp DNA linear MAM 26-MAY-1999
DEFINITION Hippopotamus amphibius gene for trappin, partial cds.
ACCESSION AB003284
VERSION AB003284.1 GI:4887643
KEYWORDS trappin.
SOURCE Hippopotamus amphibius DNA.
ORGANISM Hippopotamus amphibius
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Cetartiodactyla; Hippopotamidae; Hippopotamus.
REFERENCE
1 (sites)
AUTHORS Furutani,Y., Kato,A., Yasue,H., Alexander,L.J., Beattie,C.W. and

Db	71	TTGTCCAGGTATCAAAAAGTGCTGTGAAGGTTCTGCGGTATGGCTTGTTTCGTTCCACA	12
Qy	200	ATAATAG 206	
Db	11	ATAATAG 5	
RESULT 34			
PIGSPAI2S2			
LOCUS	PIGSPAI2S2		
DEFINITION	Porcine DNA for SPAI-2, exon 2.	483 bp	DNA
ACCESSION	D17755		
VERSION	D17755.1	GI:1054614	
KEYWORDS	SPAI-2.		
SEGMENT	2 of 3		
SOURCE	Sus scrofa cell_line NM538 (library: lambda EMBL3 SP6/T7) DNA, clone lamdaPSG1.		
ORGANISM	Sus scrofa		
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.		
AUTHORS	1 (bases 1 to 483) Kuroki,J., Hosoya,T., Itakura,M., Hirose,S., Tamechika,I., Yoshimoto,T., Ghoneim,M.A., Nara,K., Kato,A., Suzuki,Y., Furukawa,M. and Tachibana,S.		
TITLE	Cloning, characterization, and tissue distribution of porcine SPAI, a protein with a transglutaminase substrate domain and the WAP motif		
JOURNAL	J. Biol. Chem. 270 (38), 22428-22433 (1995)		
MEDLINE	95403443		
REFERENCE	2 (bases 1 to 483)		
AUTHORS	Hirose,S.		
TITLE	Direct Submission		
JOURNAL	Submitted (27-SEP-1993) Shigehisa Hirose, Tokyo Institute of Technology, Department of Biological Sciences; 4259 Nagatsuta-cho, Midori-ku, Yokohama, Kanagawa 226-8501, Japan (E-mail:shirose@bio.titech.ac.jp, Tel:045-924-5726, Fax:045-924-5824)		
COMMENT	Submitted (27-Sep-1993) to DDBJ by: Shigehisa Hirose Department of Biological Sciences Tokyo Institute of Technology 4259 Nagatsuta-cho, Midori-ku		

045-924-5726
 Phone: 045-924-5824.
 Fax:

Location/Qualifiers
 1. .483
 /organism="Sus scrofa"
 /db_xref="taxon:9823"
 /cell_line="NM538"
 /clone_lib="lambdaEMBL3 SP6/T7"
 join(DI7754.1:1. .82,1. .482)
 /note="preproSPAI-2"
 /codon_start=1
 /product="SPAI-2"
 /protein_id="BAA04603.1"
 /db_xref="GI:1054617"
 /translation="MRSRSFLVLVAVFLICETLVAQRDLRIRGPKGQGDPEVGQDQD
 EGPGPVKEILDIGQDPVKGDPEVGQDPEVGQDPEVGQDPEVGQDPEVGQDPEVGQDPEVGQD
 VYKGHEPVEGQDPVNAQLPDKVQDPVKAQPAVPGRFLLSKRGHCPRIILFRCPILSNPSN
 KWRDYDCPGVKKCEGFCGKDCLYPK"
 1. .483
 /number=2

BASE COUNT 144 a 111 c 123 g 105 t
 ORIGIN

Query Match 32.1%; Score 66.2; DB 4; Length 483;
 Best Local Similarity 66.4%; Pred. No. 8e-10;
 Matches 95; Conservative 0; Mismatches 48; Indels 0; Gaps 0;

Ov 61 CTAAGCCAGGTTCTGTCTATTATCTTGATTCGTCGCTATGTTAAACCCACCTAAC 120

```

Db      548 AGGATTGTTTGTATCCCAAGTGA 570

RESULT 36
DMY245 DMY245          789 bp      mRNA       linear     MAM 01-FEB-2000
LOCUS   Sus scrofa mRNA for preproSPAI-2, complete cds.
DEFINITION
ACCESSION D83667.1 GI:1304046
VERSION   APAl-2; preproSPAI-2; proSPAI-2; SPAI-2.
KEYWORDS  Sus scrofa cell_line:NM514 cDNA to mRNA, clone_lib:lambda gt10.
SOURCE    Sus scrofa
ORGANISM  Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
          Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
REFERENCE 1 (sites)
AUTHORS   Kuroki,J., Hosoya,T., Itakura,M., Hirose,S., Tamechika,I.,
          Yoshimoto,T., Ghoneim,M.A., Nara,K., Kato,A., Suzuki,Y.,
          Furukawa,M. and Tachibana,S.
TITLE     Cloning, characterization, and tissue distribution of porcine SPAl
          a protein with a transglutaminase substrate domain and the WAP
          motif
JOURNAL   J. Biol. Chem. 270 (38), 22428-22433 (1995)
MEDLINE   95403443
REFERENCE 2 (bases 1 to 789)
AUTHORS   Hirose,S.
TITLE     Direct Submission
JOURNAL   Submitted (23-FEB-1996) Shigehisa Hirose, Tokyo Institute of
          Technology, Department of Biological Sciences; Okayama, Meguro-ku,
          Tokyo 152, Japan (Tel:81-3-3726-1111, Fax:81-3-3729-0335)
COMMENT   On or before Mar 17, 1999 this sequence version replaced
          gi:2078450, gi:1054611, gi:2077950.
DI1753:Submitted (27-Sep-1993) to DDBJ by:Shigehisa Hirose.

FEATURES             Location/Qualifiers
     source            1..789
                     /organism="Sus scrofa"
                     /db_xref="taxon:9823"
                     /cell_line="NM514"
                     /clone_lib="lambda gt10"
                     gl..624
     CDS               1..624
                     /codon_start=1
                     /product="preproSPAI-2"
                     /protein_id="BAA12037.1"
                     /db_xref="GI:1304181"
                     /translation="MRSRSFLVLVAVFLICETLVAQLDRIRPKGGQDPVEGQDQ
EGPVPKVEILDIGQDPVKGDPAVKGDPAVKGDPAVKGDPAVKGDPAVKGDPAVKAEIPDIGD
VVGHEPVEGQDPVNAQLPDKVQDPVKAQPAVGRFLLSKRGHCPRILFRCLSLNPSN
KCWRDYDCPGVKKCCCGFCGCKDLYPK"
sig_peptide         61..123
mat_peptide         124..621
                     /product="proSPAI-2"
                     439..621
mat_peptide         /product="SPAI-2"
polyA_site          722
polyA_signal        761..766
BASE COUNT 201 a 215 c 197 g 176 t
ORIGIN

Query Match      32.1%; Score 66.2; DB 4; Length 789;
Best Local Similarity 66.4%; Pred. No. 7.9e-10;
Matches 95; Conservative 0; Mismatches 48; Indels 0; Gaps 0;

QY  61 CTAGCCAGGTTCTTGTCCTATTATCTTGATTCGTTGCGCTATGTTAAACCCACCCTAAC 120
||||| ||| ||||| ||| ||||| ||| ||||| ||| ||||| ||| ||||| |||
Db  482 CTAGCGTGGCCACTGCCCTAGGATTCITTTTCGTTGCCCGCTGAGCAATCCTCTACA 541
||||| ||| ||||| ||| ||||| ||| ||||| ||| ||||| ||| ||||| |||
QY  121 GTTGTTTGAGGACACTGATTGTCCAGGTATCAAAAAGTGCTGTGAAGGTTCTCTGCGGTA 180
||||| ||| ||||| ||| ||||| ||| ||||| ||| ||||| ||| ||||| |||
Db  542 AGTGTGGAGAGATTATGACTGTCCAGGGTCAAGAAGTGCTGTGAAGGCTTTTGC GGGA 601
||||| ||| ||||| ||| ||||| ||| ||||| ||| ||||| ||| ||||| |||
QY  181 TGGCTTGTTCGTTCCACAATAA 203
||||| ||| ||||| ||| ||||| ||| ||||| ||| ||||| ||| ||||| |||
Db  602 AGGATTGTTGTATCCCAAGTGA 624
||||| ||| ||||| ||| ||||| ||| ||||| ||| ||||| ||| ||||| |||

```

```
RESULT 37
PIGWAPB
LOCUS
DEFINITION Pig DNA for SPAI-2, complete cds.
ACCESSION D50320
VERSION D50320.1 GI:1228059
KEYWORDS SPAI-2.
SOURCE Sus scrofa DNA, clone_lib:EMBL SP6/T7 clone:lambda WAP-2.
ORGANISM Sus scrofa
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
1 (sites)
REFERENCE Tamechika,I., Itakura,M., Saruta,Y., Furukawa,M., Kato,A.,
AUTHORS Tachibana,S. and Hirose,S.
TITLE Accelerated evolution in inhibitor domains of porcine elafin family
members
JOURNAL J. Biol. Chem. 271 (12), 7012-7018 (1996)
MEDLINE 96215132
REFERENCE 2 (bases 1 to 3782)
AUTHORS Hirose,S.
JOURNAL Unpublished
REFERENCE 3 (bases 1 to 3782)
AUTHORS Hirose,S.
TITLE Direct Submission
JOURNAL Submitted (18-APR-1995) Shigehisa Hirose, Tokyo Institute of
Technology, Department of Biological Sciences; 4259 Nagatsuta-cho,
Midori-ku, Yokohama, Kanagawa 226-8501, Japan
(E-mail:shirose@bio.titech.ac.jp, Tel:045-924-5726,
Fax:045-924-5824)
FEATURES
source
1..3782
/organism="Sus scrofa"
/db_xref="taxon:9823"
/clone="lambda WAP-2"
/clone_lib="EMBL SP6/T7"
<1..710
/number=1
/evidence=experimental
join(629..710,1552..2033)
/codon_start=1
/evidence=experimental
/product="SPAI-2"
/protein_id="BAA08855.1"
/db_xref="GI:1228060"
/translation="MRSRSFLVAVFLICETLVAQLDRIRGPKGGQDPVEGDDOD
EGGPVKVEILDIGQDPVKGQDPVKGQDPVKQDPVKQDPLVKSQDPVKAELPDICQD
VWKGHPVEGQDPVNAQLPDKQDPVKAQPAVPGRFLLSKRGHCPRILFRCLSPNSN
KCWRDYDCPGVKKCCGFCGKDCLYPK"
629..691
/evidence=experimental
1848..2030
/product="SPAI-2"
/evidence=experimental
711..1551
/number=1
/evidence=experimental
1552..2034
/number=2
/evidence=experimental
join(2034,2545..2701)
2035..2544
/number=2
/evidence=experimental
2240..2502
/number=3
/evidence=experimental
2545..2701
/number=3
/evidence=experimental
2680..2685
/number=3
/evidence=experimental
2701
/evidence=experimental
sig_peptide
mat_peptide
intron
exon
3'UTR
intron
repeat_unit
exon
polyA_signal
polyA_site
```

```
repeat_region 3671..3742
/note="TG repeat"
BASE COUNT 823 a 972 c 1149 g 838 t
ORIGIN
Query Match 32.1%; Score 66.2; DB 4; Length 3782;
Best Local Similarity 66.4%; Pred. No. 7.7e-10;
Matches 95; Conservative 0; Mismatches 48; Indels 0; Gaps 0;
Qy 61 CTAAGCCAGGTTCTTGTCCTATTATCTTGATTGCTGCGCTATGTTAAACCCACCTAAC 120
Db 1891 CTAAGCGTGGCCACTGCCCTAGGATTCTTTTCGTTGCCCGCTGAGCAATCCCTCTAAC 1950
Qy 121 GTGTTTGAAGGACACTGATTGTCCAGGTATCAAAAAGTCTGTGAAGGTTCTTCCGGTA 180
Db 1951 AGTGTGGAGAGATTATGACTGTCCAGGGGTCAAGAAGTGTGTGAAGGCTTTTGGCGGA 2010
Qy 181 TGGCTTGTTCGTTCCACAATAA 203
Db 2011 AGGATTGTTTGTATCCCAAGTGA 2033
RESULT 38
A31079/c
LOCUS A31079 66 bp DNA linear PAT 21-AUG-1995
DEFINITION oligonucleotide ELI2 from patent EP0402068.
ACCESSION A31079
VERSION A31079.1 GI:1249293
KEYWORDS synthetic construct.
SOURCE synthetic construct.
ORGANISM artificial sequences.
REFERENCE 1 (bases 1 to 66)
AUTHORS Christophers,E., Schroder,J.M., Pioli,D., Wiedow,O. and Edge,M.D.
TITLE Polypeptides and polypeptide analogues with inhibitory activity
against human elastase
JOURNAL Patent: EP 0402068-A 19 12-DEC-1990;
IMPERIAL CHEMICAL INDUSTRIES PLC
FEATURES
source
1..66
/organism="synthetic construct"
/db_xref="taxon:32630"
BASE COUNT 15 a 15 c 18 g 18 t
ORIGIN
Query Match 32.0%; Score 66; DB 6; Length 66;
Best Local Similarity 100.0%; Pred. No. 9.6e-10;
Matches 66; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 5 CGAGTCGGTACCATACCTGCATATGCTCAAGAACAGTTAAAGGTCCTGTGTCTACTAA 64
Db 66 CGAGTCGGTACCATACCTGCATATGCTCAAGAACAGTTAAAGGTCCTGTGTCTACTAA 7
Qy 65 GCCAGG 70
Db 6 GCCAGG 1
RESULT 39
A31078
LOCUS A31078 64 bp DNA linear PAT 21-AUG-1995
DEFINITION oligonucleotide ELI1 from patent EP0402068.
ACCESSION A31078
VERSION A31078.1 GI:1249292
KEYWORDS synthetic construct.
SOURCE synthetic construct.
ORGANISM artificial sequences.
REFERENCE 1 (bases 1 to 64)
AUTHORS Christophers,E., Schroder,J.M., Pioli,D., Wiedow,O. and Edge,M.D.
TITLE Polypeptides and polypeptide analogues with inhibitory activity
against human elastase
JOURNAL Patent: EP 0402068-A 18 12-DEC-1990;
```

```
FEATURES
  source
    IMPERIAL CHEMICAL INDUSTRIES PLC
    Location/Qualifiers
      1..64
        /organism="synthetic construct"
        /db_xref="taxon:32630"
BASE COUNT    19 a   16 c   12 g   17 t
ORIGIN

Query Match      31.1%; Score 64; DB 6; Length 64;
Best Local Similarity 100.0%; Pred. No. 4e-09;
Matches 64; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AATTCGAGCTCGGTACCATACCTGCATATGCTCAAGAACCAGTTAAAGGTCCTGTGTCTA 60
    |||||
Db 1 AATTCGAGCTCGGTACCATACCTGCATATGCTCAAGAACCAGTTAAAGGTCCTGTGTCTA 60
    |||||
QY 61 CTAA 64
    |||||
Db 61 CTAA 64
    |||||

RESULT 40
AB003282
LOCUS          AB003282              1325 bp      DNA          linear      MAM 26-MAY-1999
DEFINITION    Phacochoerus aethiopicus gene for SPAI (trappin-1), partial cds.
ACCESSION     AB003282
VERSION       AB003282.1 GI:4887639
KEYWORDS      SPAI (trappin-1).
SOURCE        Phacochoerus aethiopicus DNA.
ORGANISM      Phacochoerus aethiopicus
              Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
              Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Phacochoerus.
REFERENCE     1 (sites)
AUTHORS       Furutani,Y., Kato,A., Yasue,H., Alexander,L.J., Beattie,C.W. and
              Hirose,S.
              Evolution of the trappin multigene family in the Suidae
              J. Biochem. 124 (3), 491-502 (1998)
              98391820
REFERENCE     2 (sites)
AUTHORS       Furutani,Y., Kato,A., Yasue,H. and Hirose,S.
              (Very Recent) evolution in suidae of the trappin multigene family
              with unusually conserved intron sequences
              Unpublished
              3 (bases 1 to 1325)
              Hirose,S.
              Direct Submission
              Submitted (21-APR-1997) Shigehisa Hirose, Tokyo Institute of
              Technology, Department of Biological Sciences; 4259 Nagatsuta-cho,
              Midori-ku, Yokohama, Kanagawa 226-8501, Japan
              (E-mail:shiroyose@bio.titech.ac.jp, Tel:045-924-5726,
              Fax:045-924-5824)
              Location/Qualifiers
                1..1325
                  /organism="Phacochoerus aethiopicus"
                  /db_xref="taxon:85517"
                  <203..828
                  /codon_start=3
                  /product="SPAI (trappin-1)"
                  /protein_id="BAA77826.1"
                  /db_xref="GI:4887640"
                  /translation="GPKSQGQDPVEGQDHEGQPVKVEILDIGQDLVKGDQPVKQGD
                  LVKSQDPVKAELPDIGQDVVKGDQPVKQDLVKSQDPVTAELPDIGQDVVKGDQPVKG
                  QDLVKSQDPVTAELPDIGQDVVKGHEPVGQDPINAIIPDKVQDPKAQPAVQGLLFL
                  SKRGCPWILLRCLPLANPSNKCWRDYDCPGVKKCEGFCGKDCLYPK"
                203..828
              exon
BASE COUNT      363 a   319 c   367 g   276 t
ORIGIN

Query Match      30.6%; Score 63; DB 4; Length 1325;
Best Local Similarity 65.0%; Pred. No. 7.8e-09;
Matches 93; Conservative 0; Mismatches 50; Indels 0; Gaps 0;

QY 61 CTAAGCCAGGTTCTTGCTATTATCTTGATTCGTTGGCTATGTTAAACCCACCTAAC 120
```

```
Db 686 CTAAGCGTGGCGCTGCCCTTGGAATCTTCCGTTGCCGCTGGCCAATCCCTCTAACA 745
    ||||| || ||||| || ||||| || ||||| || ||||| || ||||| || |||||
QY 121 GTTGTGTTGAAGGACACTGATTGTCACAGGTATCAAAAAGTGTCTGAAGGTTCTTCGCGGTA 180
    ||||| || ||||| || ||||| || ||||| || ||||| || ||||| || |||||
Db 746 AGTGTGGAGAGATTATGACTGTCCAGGGGTCAAGAAAGTGTCTGAAGGCTTTTTCGCGGA 805
    ||||| || ||||| || ||||| || ||||| || ||||| || ||||| || |||||
QY 181 TGGCTTGTTCGTTCCACAATAA 203
    ||||| || ||||| || ||||| || ||||| || ||||| || ||||| || |||||
Db 806 AGGATTGTTGTATCCCAAGTGA 828
    ||||| || ||||| || ||||| || ||||| || ||||| || ||||| || |||||

RESULT 41
PIGWAPE
LOCUS          PIGWAPE              578 bp      mRNA          linear      MAM 14-APR-2000
DEFINITION    Pig mRNA for elafin family member protein, complete cds.
ACCESSION     D50323
VERSION       D50323.1 GI:1228065
KEYWORDS      elafin family member protein.
SOURCE        Sus scrofa cDNA to mRNA, clone:WAP-4.
ORGANISM      Sus scrofa
              Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
              Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
REFERENCE     1 (sites)
AUTHORS       Tamechika,I., Itakura,M., Saruta,Y., Furukawa,M., Kato,A.,
              Tachibana,S. and Hirose,S.
              Accelerated evolution in inhibitor domains of porcine elafin family
              members
              J. Biol. Chem. 271 (12), 7012-7018 (1996)
              96215132
REFERENCE     2 (sites)
AUTHORS       Furutani,Y., Kato,A., Yasue,H., Alexander,L.J., Beattie,C.W. and
              Hirose,S.
              Evolution of the trappin multigene family in the Suidae
              J. Biochem. 124 (3), 491-502 (1998)
              98391820
REFERENCE     3 (bases 1 to 578)
AUTHORS       Hirose,S.
              Direct Submission
              Submitted (18-APR-1995) Shigehisa Hirose, Tokyo Institute of
              Technology, Department of Biological Sciences; 4259 Nagatsuta-cho,
              Midori-ku, Yokohama, Kanagawa 226-8501, Japan
              (E-mail:shiroyose@bio.titech.ac.jp, Tel:045-924-5726,
              Fax:045-924-5824)
              Location/Qualifiers
                1..578
                  /organism="Sus scrofa"
                  /db_xref="taxon:9823"
                  /clone="WAP-4"
                  <1..96
                  /number=1
                  /note="PCR primer"
                  <1..14
                  15..560
                  /codon_start=1
                  /product="elafin family member protein"
                  /protein_id="BAA08858.1"
                  /db_xref="GI:1864016"
                  /translation="MRSRSLVAVFLICETLVAQPEKIRPKGQGPVEGQDQD
                  KGQGPVKVEILDIGQDLVKRPDPVKGDQPVKQDLVKSQDPVKAELPDIGQDVVKGHD
                  PVEGQDPVNAQLPDKVQDPVKAQPPVQGRLLHYKPLGCLPWFILRCPLPKPKNKCWRDS
                  HCPGVMKCCGFCGNECSYPR"
                15..77
                375..557
                  /product="unnamed"
                  97..561
                  /number=2
                  561..578
                  562..578
                  /number=3
                  complement(562..578)
                  /note="PCR primer"
              sig_peptide
              mat_peptide
              exon
              3'UTR
              exon
              primer_bind
```

BASE COUNT 169 a 141 c 138 g 130 t
ORIGIN

Query Match 29.3%; Score 60.4; DB 4; Length 578;
Best Local Similarity 64.1%; Pred. No. 5.1e-08;
Matches 91; Conservative 0; Mismatches 51; Indels 0; Gaps 0;

QY 62 TAAGCCAGGTTCTTGTCTTATCTTATCTTGTGCGCTATGTTAAACCCACCTAACCG 121
|||||
Db 419 TAAGCCTGGCCTCTGCCCTTGGATTTTACGTGCGCGCTGCCCAACCTCTTAACAA 478
|||||

QY 122 TTGTTTGAAGGACACTGATGTCCAGGTATCAAAAAGTCTGTGAAGTTCCTGCGGTAT 181
|||||
Db 479 ATGTTGGAGAGATTCTCACTGCCAGGGTCAATGAAGTCTGTGAAGGCTTTTTCGGGAA 538
|||||

QY 182 GGCTTGTTCGTTCCACAAATAA 203
|||||
Db 539 TGAATGTTCTGATATCCCGATGA 560
|||||

RESULT 42
AB011010
LOCUS AB011010 495 bp DNA linear MAM 14-APR-2000
DEFINITION Bos taurus gene for Trappin-6, partial cds.
ACCESSION AB011010
VERSION AB011010.1 GI:3132273
KEYWORDS Trappin-6.
SOURCE Bos taurus DNA.
ORGANISM Bos taurus

REFERENCE 1 (sites)
AUTHORS Furutani,Y., Kato,A., Yasue,H., Alexander,L.J., Beattie,C.W. and Hirose,S.
TITLE Evolution of the trappin multigene family in the Suidae
JOURNAL J. Biochem. 124 (3), 491-502 (1998)
MEDLINE 98391820
AUTHORS Hirose,S.

REFERENCE 2 (bases 1 to 495)
AUTHORS Hirose,S.
TITLE Direct Submission
JOURNAL Submitted (10-FEB-1998) Shigehisa Hirose, Tokyo Institute of Technology, Department of Biological Sciences; 4259 Nagatsuta-cho, Midori-ku, Yokohama, Kanagawa 226-8501, Japan
(E-mail:shirose@bio.titech.ac.jp, Tel:+81-45-924-5726, Fax:+81-45-924-5824)

FEATURES
source Location/Qualifiers
1..495
/organism="Bos taurus"
/db_xref="taxon:9913"
intron <1..36
/number=1
primer_bind 1..24
/note="PCR primer site"
CDS <37..329
/codon_start=3
/product="Trappin-6"
/protein_id="BAA28148.1"
/db_xref="GI:3132274"
/translation="SPKQGNVFNKGPVNGQSPDKQDPVKQDPVKQDVVVVAQD
RAGLPFRGLCPVRVIRHCLNPPNQWRDAHCPGAKKCEGFCGKTCMNPR"

exon 37..>329
primer_bind 472..495
/note="PCR primer site"
BASE COUNT 134 a 99 c 163 g 99 t
ORIGIN

Query Match 26.5%; Score 54.6; DB 4; Length 495;
Best Local Similarity 61.7%; Pred. No. 3.3e-06;
Matches 87; Conservative 0; Mismatches 54; Indels 0; Gaps 0;

QY 63 AAGCCAGGTTCTTGTCTTATCTTATCTTGTGCGCTATGTTAAACCCACCTAACCGT 122
|||||

Db 189 AAGCGTGGCTTATGCCCCAGGGTTCGGATCCCACTGCAACTTGTGGAATCCCCCTAACCAAG 248
QY 123 TGTTTGAAGGACACTGATGTCCAGGTATCAAAAAGTCTGTGAAGTTCCTGCGGTATG 182
|||||
Db 249 TGTGGAGAGATGCTCACTGCCAGGGGCCCAAGAAAGTCTGTGAAGGCTTTTGTGGGAAG 308
|||||
QY 183 GCTTGTTCGTTCCACAAATAA 203
|||||
Db 309 ACCTGATGAATCCCGGTGA 329
|||||

RESULT 43
PIGWAPC
LOCUS PIGWAPC 3670 bp DNA linear MAM 14-APR-2000
DEFINITION Pig DNA for elafin family member protein, complete cds.
ACCESSION D50321
VERSION D50321.1 GI:1228061
KEYWORDS elafin family member protein.
SOURCE Sus scrofa DNA, clone_lib:EMBL SP6/T7 clone:lambda WAP-3.
ORGANISM Sus scrofa

REFERENCE 1 (sites)
AUTHORS Tamechika,I., Itakura,M., Saruta,Y., Furukawa,M., Kato,A., Tachibana,S. and Hirose,S.
TITLE Accelerated evolution in inhibitor domains of porcine elafin family members
JOURNAL J. Biol. Chem. 271 (12), 7012-7018 (1996)
MEDLINE 96215132
REFERENCE 2 (bases 1 to 3670)
AUTHORS Hirose,S.

JOURNAL Unpublished
REFERENCE 3 (bases 1 to 3670)
AUTHORS Hirose,S.
TITLE Direct Submission
JOURNAL Submitted (18-APR-1995) Shigehisa Hirose, Tokyo Institute of Technology, Department of Biological Sciences; 4259 Nagatsuta-cho, Midori-ku, Yokohama, Kanagawa 226-8501, Japan
(E-mail:shirose@bio.titech.ac.jp, Tel:045-924-5726, Fax:045-924-5824)

FEATURES
source Location/Qualifiers
1..3670
/organism="Sus scrofa"
/db_xref="taxon:9823"
/clone="lambda WAP-3"
/clone_lib="EMBL SP6/T7"
exon <1..710
/number=1
5'UTR <1..628
CDS join(629..710,1552..1904)
/codon_start=1
/product="elafin family member protein"
/protein_id="BAA08856.1"
/db_xref="GI:1228062"
/translation="MRSRSFLVLVAVFLICETLVAQRLDRIRGPKGGQDPVEGQDQD
EGQGPVKVEILDIGQDLVKQDPVEGQDPVKAQLPDKVQDPVKAQPIQGGFLFPKPG
VCPKIIIFCPLVNPPIKCWRDSCHPGVKKCCPSLCGKGCVTPR"

sig_peptide 629..691
mat_peptide 1722..1901
intron /product="unnamed"
711..1551
/number=1
exon 1552..1905
/number=2
3'UTR join(1905,2414..2570)
intron 1906..2413
/number=2
repeat_unit 2110..2371
/note="SINE (PRE-1)"
exon 2414..2570
/number=3
polyA_signal 2549..2554
polyA_site 2570

QY 181 TGGCTTGTTCGTTCCACAATAATA 205

REFERENCE	2	Furutani, Y. and Hirose, S.
AUTHORS		
TITLE		Evolution of caltrin-like protein
JOURNAL		Unpublished
REFERENCE	3	(bases 1 to 464)
AUTHORS		Furutani, Y. and Hirose, S.
TITLE		Direct Submission
JOURNAL		Submitted (26-APR-2000) Shigehisa Hirose, Tokyo Institute of Technology, Department of Bioscience; Nagatudacho4259 Midori-ku, Yokohama, Kanagawa 226-8501, Japan (E-mail: shirose@bio.titech.ac.jp, Tel: 81-45-924-5726, Fax: 81-45-924-5824)
FEATURES		Location/Qualifiers
source	1.	.464

```
/organism="Cavia porcellus"
/db_xref="taxon:10141"
/tissue_lib="seminal vesicle"
17. .247
/notes="antileukoproteinase repeat homology #label ALP"
/codon_start=1
/product="caltrin-like protein II"
/protein_id="BAB70710.1"
/db_xref="GI:16519355"
/translation="MKATILFILELLEKPSFGRRLLHGQAINRPGSCPRVMIVCPA
RHPPNKTSDYDCPKPKQCCPGYCGKQCYQE"
17. .79
sig_peptide
mat_peptide

misc_feature      113 a      139 c      97 g      115 t
/notes="WAP. domain"

BASE COUNT      113 a      139 c      97 g      115 t
ORIGIN
Query Match      18.6%; Score 38.4; DB 10; Length 464;
Best Local Similarity 54.2%; Pred. No. 0.36;
Matches 78; Conservative 0; Mismatches 66; Indels 0; Gaps 0;

Qy 60 ACTAAGCCAGGTTCTTGTCTATTATCTTGTGCTGCGCTATGTTAAACCCACCTAAC 119
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 104 AATAGGCTGTTCTGCTGCGGGTTATGATATATACTGCGCGCAAGGCATCTCCCAAC 163

Qy 120 CGTTGTTGAAGGACACTGATGTCTCCAGGTATCAAAAAGTGTGTGAAGGTTCTCTCGGTT 179
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 164 AAGTGACGAGTGATTATGACTGCCCAAGCCTCAGAAGTGTGCCCAGGCTACTGCGGG 223

Qy 180 ATGGCTTGTTCTGTTCCACATAA 203
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 224 AAGCAATGTTATCAACCTGAATGA 247

RESULT 47
AC125967
LOCUS
DEFINITION
Rattus norvegicus clone CH230-170K14, *** SEQUENCING IN PROGRESS
***, 29 unordered pieces.
AC125967
AC125967.3 GI:21956531
VERSION
HTG; HTGS PHASE1.
KEYWORDS
Norway rat.
SOURCE
Rattus norvegicus
ORGANISM
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
Rattus.
REFERENCE
1 (bases 1 to 204225)
Muzny,D.M., Adams,C., Adio-Oduola,B., Ali-osman,F.R., Allen,C.,
Alsbrooks,S.L., Amaratunge,H.C., Are,J.R., Ayele,M., Banks,T.,
Barbaria,J., Benton,J., Bimage,K., Blankenburg,K., Bonnin,D.,
Bouck,J., Bowie,S., Brieva,M., Brown,E., Brown,M., Bryant,N.P.,
Buhay,C., Burch,P., Burkett,C., Burrell,K.L., Byrd,N.C.,
Carron,T.F., Carter,M., Cavazos,S.R., Chacko,J., Chavez,D.,
Chen,G., Chen,R., Chen,Z., Chowdhry,I., Christopoulos,C.,
Cleveland,C.D., Cox,C., Coyle,M.D., Dathorne,S.R., David,R.,
Davila,M.L., Davis,C., Davy-Carroll,L., Dederich,D.A.,
DeLaney,K.R., Delgado,O., Denn,A.L., Ding,Y., Dinh,H.H.,
Douthwaite,K.J., Draper,H., Dugan-Rocha,S., Durbin,K.J.,
Earhart,C., Edgar,D., Edwards,C.C., Elhaj,C., Escotto,M.,
Falls,T., Ferraguto,D., Flagg,N., Ford,J., Foster,P., Frantz,P.,
Gabisi,A., Gao,J., Garcia,A., Garner,T., Garza,N., Gill,R.,
Gorrell,J.H., Guevara,W., Gunaratne,P., Hale,S., Hamilton,K.,
Hernandez,O., Hodgson,A., Hogues,M., Holloway,C., Hollins,B.,
Homsí,F., Howard,S., Huber,J., Hulyk,S., Hume,J., Jackson,L.E.,
Jacobson,B., Jia,Y., Johnson,R., Jolivet,S., Joudah,S.,
Karlsson,E., Kelly,S., Khan,U., King,L., Korvah,J., Kovar,C.,
Kratovic,J., Kureshi,A., Landry,N., Leal,B., Lewis,L.C., Lewis,L.,
Li,J., Li,Z., Lichtarge,O., Lieu,C., Liu,J., Liu,W., Loulseged,H.,
```

```
Lozado,R.J., Lu,X., Lucier,A., Lucier,R., Luna,R., Ma,J.,
Maheshwari,M., Mapua,P., Martin,R., Martindale,A., Martinez,E.,
Massey,E., Mawhney,E., McLeod,M.P., Meador,M., Mei,G., Metzker,M.,
Miner,G., Miner,Z., Mitchell,T., Mohabbat,K., Morgan,M., Morris,S.,
Moser,M., Neal,D., Newton,J., Newton,N., Nguyen,A., Nguyen,N.,
Nguyen,N., Nickerson,E., Nwokenkwo,S., Oguh,M., Okwuonu,G.,
Oragunye,N., Oviedo,R., Pace,A., Payton,B., Peery,J., Perez,L.,
Peters,L., Pickens,R., Primus,E., Pu,L.L., Quiles,M., Ren,Y.,
Rives,M., Rojas,A., Rojubokan,I., Rolfe,M., Ruiz,S., Savery,G.,
Scherer,S., Scott,G., Shen,H., Shooshtari,N., Sisson,I.,
Sodergren,E., Sonaike,T., Sparks,A., Stanley,H., Stone,H.,
Sutton,A., Svatek,A., Tabor,P., Tamerisa,A., Tamerisa,K., Tang,H.,
Tansey,J., Taylor,C., Taylor,T., Telford,B., Thomas,N., Thomas,S.,
Usmani,K., Vasquez,L., Vera,V., Villalon,D., Vinson,R., Wang,Q.,
Wang,S., Ward-Moore,S., Warren,R., Washington,C., Watlington,S.,
Williams,G., Williamson,A., Wleczyk,R., Wooden,S., Worley,K.,
Wu,C., Wu,Y., Wu,Y.F., Zhou,J., Zorrilla,S., Nelson,D.,
Weinstock,G. and Gibbs,R.
Direct Submission
Unpublished
2 (bases 1 to 204225)
Worley,K.C.
Direct Submission
Submitted (02-JUL-2002) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA
3 (bases 1 to 204225)
Worley,K.C.
Direct Submission
Submitted (31-JUL-2002) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA
On Jul 26, 2002 this sequence version replaced gi:21954968.
----- Genome Center
Center: Baylor College of Medicine
Center code: BCM
Web site: http://www.hgsc.bcm.tmc.edu/
Contact: hgsc-help@bcm.tmc.edu
----- Project Information
Center project name: GMIB
Center clone name: CH230-170K14
----- Summary Statistics
Sequencing vector: Plasmid;
Chemistry: Dye-terminator Big Dye; 100% of reads
Assembly program: Phrap; version 0.990329
Consensus quality: 180778 bases at least Q40
Consensus quality: 184764 bases at least Q30
Consensus quality: 187452 bases at least Q20
-----
* NOTE: Estimated insert size may differ from sequence length
* (see http://www.hgsc.bcm.tmc.edu/docs/Genbank_draft_data.html).
* NOTE: This is a 'working draft' sequence. It currently
* consists of 29 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
* 1 1273: contig of 1273 bp in length
* 1274 1373: gap of unknown length
* 1374 3021: contig of 1648 bp in length
* 3022 3121: gap of unknown length
* 3122 4764: contig of 1643 bp in length
* 4765 4864: gap of unknown length
* 4865 5869: contig of 1005 bp in length
* 5870 8064: contig of 2095 bp in length
* 8065 8165: gap of unknown length
* 8165 11217: contig of 3052 bp in length
* 11217 11316: gap of unknown length
* 11317 14508: contig of 3192 bp in length
* 14509 14608: gap of unknown length
```

*	14609	18981: contig of 4373 bp in length
*	18982	19081: gap of unknown length
*	19082	22382: contig of 3301 bp in length
*	22383	22482: gap of unknown length
*	22483	24761: contig of 2279 bp in length
*	24762	24861: gap of unknown length
*	24862	28297: contig of 3436 bp in length
*	28298	28397: gap of unknown length
*	28398	31802: contig of 3405 bp in length
*	31803	31902: gap of unknown length
*	31903	37457: contig of 5555 bp in length
*	37458	37557: gap of unknown length
*	37558	43403: contig of 5846 bp in length
*	43404	43503: gap of unknown length
*	43504	49186: contig of 5683 bp in length
*	49187	49286: gap of unknown length
*	49287	56048: contig of 6762 bp in length
*	56049	56148: gap of unknown length
*	56149	63736: contig of 7588 bp in length
*	63737	63836: gap of unknown length
*	63837	69493: contig of 5657 bp in length
*	69494	69593: gap of unknown length
*	69594	77974: contig of 8381 bp in length
*	77975	78074: gap of unknown length
*	78075	83932: contig of 5858 bp in length
*	83933	84032: gap of unknown length
*	84033	92690: contig of 8658 bp in length
*	92691	92790: gap of unknown length
*	92791	103073: contig of 10283 bp in length
*	103074	103173: gap of unknown length
*	103174	114922: contig of 11749 bp in length
*	114923	115022: gap of unknown length
*	115023	127345: contig of 12323 bp in length
*	127346	127445: gap of unknown length
*	127446	138012: contig of 10567 bp in length
*	138013	138112: gap of unknown length
*	138113	148920: contig of 10808 bp in length
*	148921	149020: gap of unknown length
*	149021	164577: contig of 15557 bp in length
*	164578	164677: gap of unknown length
*	164678	180799: contig of 16122 bp in length
*	180800	180899: gap of unknown length
*	180900	204225: contig of 23326 bp in length.

BASE COUNT	ORIGIN	61390 a	37556 c	38806 g	58187 t	8286 others

	Query Match	17.3%;	Score 35.6;	DB 2;	Length 204225;
	Best Local Similarity	54.6%;	Pred. No. 2.4;		
Matches	71; Conservative	0;	Mismatches 59;	Indels 0;	Gaps 0;
QY 76	GTCCATATTATCTTGATTCGTTGGCGCTATGTTAAACCCACCCTAACCGTTGTTTGAAGGACA	135			
Db 109312	GTGCTTGTTCTTCTCTCCTCAAACCATGGTGAACTCTCCTCAATCCCTTGAGTCTGTCC	109371			
QY 136	CTGATTGTCCAGGTATCAAAAAGTGCTGTGAAGTTCCTGCGGTATGGCTTGTTCGTTTC	195			
Db 109372	CCTTTCCTCCAGGACCTAAAAAGTTCGCCCTATTCTCTGCCCTATGGCTTCTTTACTGA	109431			
QY 196	CACAAATAAATA	205			
Db 109432	CAGACCAAGA	109441			

RESULT 48	AC006054/c	AC006054	143738 bp	DNA	linear	PRI 07-MAR-2002
LOCUS		Homo sapiens	Xq28	BAC G51-382P7	(Roswell	
DEFINITION		Human BAC Library)	complete sequence.		park Cancer Institute	

ACCESSION VERSION KEYWORDS SOURCE ORGANISM	REFERENCE AUTHORS	TITLE JOURNAL REFERENCE AUTHORS TITLE JOURNAL	REFERENCE AUTHORS TITLE JOURNAL	REFERENCE AUTHORS TITLE JOURNAL	REFERENCE AUTHORS TITLE JOURNAL	COMMENT
--	----------------------	--	--	--	--	---------

AC006054
AC006054.2 GI:4557048
HTG.
Homo sapiens.
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 143738)
Muzny,D., Arenson,A.D., Bouck,J., Brundage,E., Bunac,C., Chen,Z.,
Di,W., Ding,Y., Dugan,S., Durbin,J., Forcum,J., Garcia,C.,
Gorell,J.H., Gorrell,L.L., Hernandez,J., Jackson,L.,
Kondejewski,N., Leal,B., Lichtarge,O., Liu,W., Logan,O., Lu,J.,
Martinez,C., Oswal,G., Pampell,L.R., Parish,B.J., Perez,L.,
Rashid,N.D., Rives,C., Scherer,S.E., Shen,H., Simon,M., Vo,Q.,
Williamson,A., Worley,K.C., Yu,W., Zhou,X., Nelson,D. and
Gibbs,R.A.
Direct Submission
Unpublished
2 (bases 1 to 143738)
Worley,K.C.
Direct Submission
Submitted (26-NOV-1998) Molecular and Human Genetics, Baylor
College of Medicine, One Baylor Plaza, Houston, TX 77030, USA
3 (bases 1 to 143738)
Worley,K.C.
Direct Submission
Submitted (31-MAR-1999) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA
4 (bases 1 to 143738)
Worley,K.C.
Direct Submission
Submitted (01-APR-1999) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA
5 (bases 1 to 143738)
Worley,K.C.
Direct Submission
Submitted (28-MAR-2000) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA
6 (bases 1 to 143738)
Worley,K.C.
Direct Submission
Submitted (07-MAR-2002) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA
On Mar 31, 1999 this sequence version replaced gi:4225895.
INFORMATION: <http://www.hgsc.bcm.tmc.edu/> or email
gc-help@bcm.tmc.edu

CLONE LENGTH: This sequence does not necessarily represent the entire insert of this clone. Overlapping regions of clones are only sequenced and submitted once, so the sequence for the remainder of the insert may be found in the record for the adjacent clones. Overlapping clones are noted at the beginning and end of the Features listing.

ANNOTATION OF FEATURES:

STs are identified using ePCR (Genome Res. 7:541-550) searches of a local database that includes entries from dbSTS, GDB, and local mapping efforts.

Repeats are identified using RepeatMasker (A. Smit and P. Green, unpublished.) for Human and Mouse sequences.

unpublished) for human and mouse sequences. Genes and Region of sequence similarity are identified by BLAST (Nuc. Acids Res. 25:3389-3402) similarity (expect < 1e-34) to the EST and cDNA sequences. Genes demonstrate at least two exons flanked by consensus splice sites that maintained sequence continuity across the splice junctions. Sequences that are not identical matches are annotated as similar.

SEQUENCING READ COVERAGE:Sequencing is completed to a minimum standard of double strand coverage with a minimum of 2 clones and 2

reads with no ambiguities or 2 chemistries with a minimum of 2 clones and 3 reads with no ambiguities. If the sequence quality for a region does not meet this standard, it will be indicated in the annotation as Low Coverage.

QUALITY OF INDIVIDUAL BASES: This sequence meets stringent quality standards - estimated error rate less than 1 per 10,000 bases. Reports of lowest quality individual bases and measures of base quality are listed below. Description of the metrics can be found at URL: <http://gc.bcm.tmc.edu:8088/quality.info/genbank.annotation.html>.

QUALSTAT-REPORT-----

----- Summary Statistics -----
Contig length: 143738
Phrap values in estimate: 142753
Average error rate (BCM-Phrap estimate): 0.000121373
Fraction of Phrap values less than 40 : 0.0449448
Number of consensus changing edits: 3
Number of N's in consensus : 0

----- Consensus changing edits -----
Position Original+Context Edited+Context
765 acttggtgca(n)ggaataatt acttggtgca(a)ggaataatt
28154 atttctctga(n)gatactgat atttctctga(t)gatactgat
114564 ttaaaccaaa(n)aaaacttca ttaaaccaaa(t)aaaacttca

----- Distribution of Quality < 40 Bases -----

#	bases	5	10	15	20	25	30	35	40
1000		*	*	*	*	*	*	*	*
900		*	*	*	*	*	*	*	*
800		*	*	*	*	*	*	*	*
700		*	*	*	*	*	*	*	*
600		*	*	*	*	*	*	*	*
500		*	*	*	*	*	*	*	*
400		*	*	*	*	*	*	*	*
300		*	*	*	*	*	*	*	*
200		*	*	*	*	*	*	*	*
100		*	*	*	*	*	*	*	*
0		*	*	*	*	*	*	*	*

Version: 1.01 qxfo.
FEATURES
source Location/Qualifiers
1.143738
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="X"
/clone="RPC111-382P7"
complement(1..431)
/rpt_family="L1P"
complement(2833..2856)
/rpt_family="L1P5"
complement(2863..3652)
/rpt_family="L1P5"
complement(3960..3996)
/rpt_family="AT rich"
complement(4994..5058)
/rpt_family="L1MA8"
complement(5834..5871)
/rpt_family="L1P5"
7863..7909
/rpt_family="L2"
complement(8087..8354)
/rpt_family="MIR"

repeat_region complement(9094..9263)
repeat_region /rpt_family="MIR"
repeat_region 9280..9368
repeat_region /rpt_family="L2"
repeat_region complement(9668..9704)
repeat_region /rpt_family="AT rich"
repeat_region 10431..10863
repeat_region /rpt_family="L2"
repeat_region 11375..11451
repeat_region /rpt_family="MLTIE"
repeat_region 11654..11937
repeat_region /rpt_family="MLTIE"
repeat_region 11992..12097
repeat_region /rpt_family="MIR"
repeat_region complement(12172..12831)
repeat_region /rpt_family="L2"
repeat_region 14636..14789
repeat_region /rpt_family="L2"
repeat_region 15683..15839
repeat_region /rpt_family="MIR"
repeat_region complement(15876..15996)
repeat_region /rpt_family="MERSA"
repeat_region complement(16058..16330)
repeat_region /rpt_family="L1MA4"
repeat_region complement(18548..18670)
repeat_region /rpt_family="L2"
repeat_region complement(18714..18952)
repeat_region /rpt_family="MLTIE"
repeat_region complement(18969..19123)
repeat_region /rpt_family="MLTIE"
repeat_region complement(19638..19678)
repeat_region /rpt_family="L2"
repeat_region complement(21891..21949)
repeat_region /rpt_family="MIR"
repeat_region 22022..22086
repeat_region /rpt_family="(CATG)n"
repeat_region 24250..24335
repeat_region /rpt_family="L1MA9"
repeat_region complement(24336..24395)
repeat_region /rpt_family="(TA)n"
repeat_region 24396..24906
repeat_region /rpt_family="L1MA9"
repeat_region complement(24980..25220)
repeat_region /rpt_family="MER33"
repeat_region complement(25229..25946)
repeat_region /rpt_family="L1PA10"
repeat_region complement(25944..26730)
repeat_region /rpt_family="L1PA7"
repeat_region 26728..28079
repeat_region /rpt_family="L1PA7"
misc_feature 27846..27960

Query Match 17.2%; Score 35.4; DB 9; Length 143738;
Best Local Similarity 66.2%; Pred. No. 2.8;
Matches 51; Conservative 0; Mismatches 26; Indels 0; Gaps 0;

Qy 18 ATACCTGCATATGCTCAGAACCAAGTTAAAGGTCCTGTCTACTAAGCCAGGTTCTTGT 77
Db 44303 ATGTCTCCATTTTAAAGAACCAATGCAAGTCTAATGAGTACTAGTCTAGTCTTGGT 44244
Qy 78 CCTATTATCTTGATTTCG 94
Db 44243 CCTATTATCTTGATTTCG 44227

RESULT 49
AC096647
LOCUS AC096647 146935 bp DNA linear PRI 01-MAR-2002
DEFINITION Homo sapiens BAC clone RP11-2G16 from 2, complete sequence.
ACCESSION AC096647 AC021047
VERSION AC096647.1 GI:15668086
KEYWORDS HTG.
SOURCE Homo sapiens.

```
ORGANISM      Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE     1 (bases 1 to 146935)
AUTHORS       Sulston, J.E. and Waterston, R.
TITLE         Toward a complete human genome sequence
JOURNAL       Genome Res. 8 (11), 1097-1108 (1998)
MEDLINE       99063792
PUBMED        9847074
REFERENCE     2 (bases 1 to 146935)
AUTHORS       Nguyen, C., Abbott, A., Dignan, G., Doebber, A. and Boyer, E.
TITLE         The sequence of Homo sapiens BAC clone RP11-2G16
JOURNAL       Unpublished (2001)
REFERENCE     3 (bases 1 to 146935)
AUTHORS       Waterston, R.H.
TITLE         Direct Submission
JOURNAL       Submitted (19-SEP-2001) Genome Sequencing Center, Washington
              University School of Medicine, 4444 Forest Park Parkway, St. Louis,
              MO 63108, USA
REFERENCE     4 (bases 1 to 146935)
AUTHORS       Waterston, R.
TITLE         Direct Submission
JOURNAL       Submitted (01-MAR-2002) Department of Genetics, Washington
              University, 4444 Forest Park Avenue, St. Louis, Missouri 63108, USA
              On Sep 19, 2001 this sequence version replaced gi:8072447.
              ----- Genome Center
              Center: Washington University Genome Sequencing Center
              Center code: WUGSC
              Web site: http://genome.wustl.edu/gsc
              Contact: sapiens@watson.wustl.edu
              ----- Summary Statistics
              Center project name: H_NH0002G16
              Drafting Center: WIBR
              -----

NOTICE: This sequence may not represent the entire insert of this
clone. It may be shorter because we only sequence overlapping
clone sections once, or longer because we provide a small overlap
between neighboring data submissions.

This sequence was finished as follows unless otherwise noted:
all regions were double stranded, sequenced with an alternate
chemistry, or covered by high quality data (i.e., phred quality >=
30); an attempt was made to resolve all sequencing problems, such
as compressions and repeats; all regions were covered by sequence
from more than one subclone; and the assembly was confirmed by
restriction digest.

MAPPING INFORMATION:
Mapping information for this clone was provided by Dr. John D.
McPherson, Department of Genetics, Washington University, St. Louis
MO. For additional information about the map position of this
sequence, see http://genome.wustl.edu/gsc

SOURCE INFORMATION:
The RPCI-11 human BAC library was made from the blood of one male
donor, as described by Osogawa, K., Woon, P.Y., Zhao, B., Frengen, E.,
Tateno, M., Catanese, J.J. and de Jong, P.J. (1998) An improved
approach for construction of bacterial artificial chromosome
libraries. Genomics 51:1-8. The clone may be obtained either from
Research Genetics, Inc. (http://www.resgen.com) or Pieter de Jong
and coworkers at http://www.chori.org
VECTOR: pBACe3.6

NEIGHBORING SEQUENCE INFORMATION:
The clone sequenced to the right is RP11-403L5. Actual start of
this clone is at base position 1 of RP11-2G16; actual end is at
base position 146935 of RP11-2G16.

Data from AC078976 was used to finish this clone, AC021047.

The sequence of AC021047 has been incorporated into AC096647.
Location/Qualifiers

source
1. .146935
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="2"
/map="2"
/clone="RP11-2G16"
/clone_lib="RPCI-11"
5740. .5836
/rpt_family="L1"
repeat_region
7457. .7869
/rpt_family="L1"
repeat_region
7870. .7926
/rpt_family=" (TA)n"
repeat_region
8843. .8956
/rpt_family="MaLR"
repeat_region
8976. .9017
/rpt_family="MaLR"
repeat_region
9071. .9137
/rpt_family=" (TATG)n"
repeat_region
9130. .9361
/rpt_family="MaLR"
repeat_region
9387. .9434
/rpt_family=" (TTTTTA)n"
repeat_region
10242. .10413
/rpt_family="MIR"
repeat_region
11002. .11091
/rpt_family=" (TA)n"
repeat_region
11091. .11197
/rpt_family="MaLR"
repeat_region
11211. .11331
/rpt_family="Alu"
repeat_region
11337. .11558
/rpt_family="MaLR"
repeat_region
11642. .11838
/rpt_family="L1"
repeat_region
11911. .11996
/rpt_family=" (TTTA)n"
repeat_region
12025. .12303
/rpt_family="L1"
repeat_region
12304. .12608
/rpt_family="Alu"
repeat_region
12585. .12608
/rpt_family=" (A)n"
repeat_region
12609. .12805
/rpt_family="L1"
repeat_region
13638. .14237
/rpt_family="L1"
repeat_region
14918. .15744
/rpt_family="L2"
repeat_region
16220. .16348
/rpt_family="MIR"
repeat_region
16375. .16411
/rpt_family="AT_rich"
repeat_region
17655. .17682
/rpt_family=" (TTA)n"
repeat_region
17937. .17968
/rpt_family="AT_rich"
repeat_region
18112. .18297
/rpt_family="MIR"
repeat_region
18374. .18947
/rpt_family="MER2_type"
repeat_region
18953. .19627
/rpt_family="L1"
repeat_region
22195. .22263
/rpt_family="L2"
repeat_region
24382. .24418
/rpt_family=" (TG)n"
repeat_region
24883. .25000
/rpt_family="T-rich"
repeat_region
25175. .25210
/rpt_family="AT_rich"
repeat_region
25227. .25569
/rpt_family="L1"
```

FEATURES

repeat_region 25582..25877
/rpt_family="Alu"
repeat_region 25863..25922
/rpt_family="AT_rich"
repeat_region 25895..26547
/rpt_family="L1"
repeat_region 26552..26649
/rpt_family="L1"
repeat_region 26649..27752
/rpt_family="L1"
repeat_region 27759..28143
/rpt_family="MaLR"
repeat_region 28150..28424
/rpt_family="MaLR"
repeat_region 28425..28453
/rpt_family="(T)n"
repeat_region 28605..29988
/rpt_family="MaLR"
repeat_region 30035..30248
/rpt_family="Alu"
repeat_region 30260..30999
/rpt_family="L1"
repeat_region 31000..31073
/rpt_family="MaLR"
repeat_region 31079..31311
/rpt_family="Alu"
repeat_region 31330..31623
/rpt_family="ERV1"
repeat_region 31648..31712
/rpt_family="L1"
repeat_region 31720..33118
/rpt_family="L1"
repeat_region 33111..33144
/rpt_family="AT_rich"
repeat_region 33126..33921
/rpt_family="L1"
repeat_region 35401..38116
/rpt_family="L1"
repeat_region 35601..35634
/rpt_family="AT_rich"
repeat_region 37164..37213
/rpt_family="T-rich"
repeat_region 38095..38245
/rpt_family="L1"
repeat_region 38267..39514
/rpt_family="MaLR"
repeat_region 39515..39924
/rpt_family="MaLR"

Query Match 17.1%; Score 35.2; DB 9; Length 146935;
Best Local Similarity 54.7%; Pred. No. 3.2;
Matches 70; Conservative 0; Mismatches 58; Indels 0; Gaps 0;

QY 44 TAAAGGTCCTGTGCTACTAAGCCAGGTTCTTGTCTATTATCTTGTATTCGTTGCGGTAT 103
|||||
Db 31575 TAAAGATGAGCTTTTGTGCAAGGTATGCTTTTATTGTGATAGTTTGTGTGTTGCAATTT 31634
|||||
QY 104 GTTAAACCCACCTAACCGTTGTTTGAAGGACACTGATTGTCCAGGTATCAAAAAGTGCTG 163
|||||
Db 31635 TTTAAACTAAACAATAATTGTTTATGTGTACATAATGTGTAGTATCAGATCAGGGTG 31694
|||||
QY 164 TGAAGGTT 171
|||
Db 31695 TGTAGGT 31702

RESULT 50
PIGALP PIGALP
LOCUS Porcine antileukoproteinase mRNA, complete cds.
DEFINITION Porcine antileukoproteinase mRNA, complete cds.
ACCESSION M57446
VERSION M57446.1 GI:164319
KEYWORDS antileukoproteinase.

SOURCE Porcine uterus. cDNA to mRNA, clone pALP.
ORGANISM Sus scrofa
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
REFERENCE 1 (bases 1 to 600)
AUTHORS Farmer,S.J., Fliss,A.E. and Simmen,R.C.
TITLE Complementary DNA cloning and regulation of expression of the messenger RNA encoding a pregnancy-associated porcine uterine protein related to human antileukoproteinase
JOURNAL Mol. Endocrinol. 4 (8), 1095-1104 (1990)
MEDLINE 91155942
PUBMED 2293019
FEATURES
source Location/Qualifiers
1..600
/organism="Sus scrofa"
/db_xref="taxon:9823"
/clone="pALP"
/sex="female"
/tissue_type="uterus"
1..600
/gene="ALP"
44..391
/gene="ALP"
/codon_start=1
/product="antileukoproteinase"
/protein_id="AAA63446.1"
/db_xref="GI:164320"
/translation="MAPWAVEGAENALKGACPPRKIVQCLRYEKPCKTSDWQCPDKK
KCCRDTCALKCLNPVAITNPVKPGKCPVVYGGCMMLNPNHCKTDSQCLGDLKCKK
SMCGKVCLTPVKA"
polyA_site 600
BASE COUNT 146 a 156 c 151 g 147 t
ORIGIN

Query Match 16.9%; Score 34.8; DB 4; Length 600;
Best Local Similarity 53.7%; Pred. No. 4.7;
Matches 72; Conservative 0; Mismatches 62; Indels 0; Gaps 0;

QY 63 AAGCCAGGTTCTTGTCTATTATCTTGTATTCGTTGCGGTATGTTAAACCCACCTAACCGT 122
|||||
Db 245 AAGCCTGGGAAGTGTCCAGTGGTCTATGGCCAGTGTATGCTCAACCCCAATCAC 304
|||||
QY 123 TGTTTGAAGGACACTGATTGTCCAGGTATCAAAAAGTGCTGTGAAGTTCTTCGCGGTATG 182
|||||
Db 305 TGCAAGACAGACAGCCAGTGCCTGGTGACTTTAAATGCTGCAAGAGCATGTGCGGAAA 364
|||||
QY 183 GCTTGTTTCGTTCC 196
|||
Db 365 GTCTGCCTCACCCC 378

RESULT 51
MMU88093
LOCUS Mus musculus secretory leukocyte protease inhibitor mRNA, complete cds.
DEFINITION
ACCESSION U88093
VERSION U88093.1 GI:1945382
KEYWORDS
SOURCE Mus musculus.
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 671)
AUTHORS Zitnik,R.J., Zhang,J., Kashem,M.A., Kohno,T., Lyons,D.E.,
Wright,C.D., Rosen,E., Goldberg,I. and Hayday,A.C.
TITLE The cloning and characterization of a murine secretory leukocyte protease inhibitor cDNA
JOURNAL Biochem. Biophys. Res. Commun. 232 (3), 687-697 (1997)
MEDLINE 97271386
PUBMED 9126337
REFERENCE 2 (bases 1 to 671)

```
AUTHORS      Zitnik,R.J., Zhang,J., Kashem,M.A., Kohn,T., Lyons,D.E.,
Wright,C.D., Rosen,E., Goldberg,I. and Hayday,A.C.
TITLE        Direct Submission
JOURNAL      Submitted (31-JAN-1997) Internal Medicine, Yale University, 333
Cedar Street, New Haven, CT 06520, USA
FEATURES     Location/Qualifiers
source       1..671
             /organism="Mus musculus"
             /db_xref="taxon:10090"
             /cell_line="ras-transformed 3T3 cells"
             13..408
             /function="serine protease inhibitor"
             /codon_start=1
             /product="secretory leukocyte protease inhibitor"
             /protein_id="AAC53140.1"
             /db_xref="GI:1945383"
             /translation="MKSCGLLPFTVLLALGILAPWTVEGGKNDAIKIGACPAKKPAQC
             LKLEKPCRTDWECPGKQRCQDAGSKCVNPVPIRKPVWRKPGRCVKVTQARCMMLNP
             PNVQORDGQCDGKYKCCGICGKVCCLPPM"
BASE COUNT  157 a 179 c 177 g 158 t
ORIGIN
Query Match      16.8%; Score 34.6; DB 10; Length 671;
Best Local Similarity 53.3%; Pred. No. 5.4;
Matches 73; Conservative 0; Mismatches 64; Indels 0; Gaps 0;

QY  51  CCTGTGCTACTAAGCCAGGTTCTTGTCTTATCTTGTGCTTGGCTATGTTAAAC 110
      |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
Db   256  CCAGTGTGGAGGAGCTGGAGGTCGCTCAAAACTCAGGCAAGATGTATGATGCTTAAC 315

QY  111  CCACCTAACCGTTGTTTGAAGGACACTGATTGTCCAGGTATCAAAAAGTCTGTGAAGGT 170
      |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
Db   316  CCTCCCAATGTCTGCCAGAGGGACGGGCAGTGTGACGGCAAATACAAGTCTGTGAGGTT 375

QY  171  TCCTGCGGTATGGCTTG 187
      |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
Db   376  ATATGTGGGAAAGTCTG 392

RESULT 52
MMU94341
LOCUS      MMU94341      682 bp      mRNA      linear      ROD 29-OCT-1997
DEFINITION Mus musculus secretory leukoprotease inhibitor mRNA, complete cds.
ACCESSION  U94341
VERSION     U94341.1  GI:1945450
KEYWORDS
SOURCE      Mus musculus.
ORGANISM    Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
REFERENCE   1  (bases 1 to 682)
AUTHORS     Abe,T., Tominaga,Y., Kikuchi,T., Watanabe,A., Satoh,K., Watanabe,Y.
            and Nukiwa,T.
TITLE       Bacterial pneumonia causes augmented expression of the secretory
            leukoprotease inhibitor gene in the murine lung
JOURNAL     Am. J. Respir. Crit. Care Med. 156 (4 Pt 1), 1235-1240 (1997)
MEDLINE     98011992
PUBMED      9351627
REFERENCE   2  (bases 1 to 682)
AUTHORS     Abe,T., Tominaga,Y., Kikuchi,T., Watanabe,A., Satoh,K., Watanabe,Y.
            and Nukiwa,T.
TITLE       Direct Submission
JOURNAL     Submitted (18-MAR-1997) Department of Respiratory Oncology and
            Molecular Medicine, Institute of Development, Aging and Cancer,
            Tohoku University, 4-1 Seiryomachi, Aoba-ku, Sendai 980-77, Japan
FEATURES     Location/Qualifiers
source       1..682
             /organism="Mus musculus"
             /strain="B6CBA [C57Black/6 x CBA]"
             /db_xref="taxon:10090"
             /tissue type="lung"
             14..409
             /codon_start=1

CDS
            1..409
```

```
/product="secretory leukoprotease inhibitor"
/protein_id="AAC53394.1"
/db_xref="GI:1945451"
/translation="MKSCGLLPFTVLLALGILAPWTVEGGKNDAIKIGACPAKKPAQC
LKLEKPCRTDWECPGKQRCQDAGSKCVNPVPIRKPVWRKPGRCVKVTQARCMMLNP
PNVCQORDGQCDGKYKCCGICGKVCCLPPM"
BASE COUNT  163 a 179 c 181 g 159 t
ORIGIN
Query Match      16.8%; Score 34.6; DB 10; Length 682;
Best Local Similarity 53.3%; Pred. No. 5.4;
Matches 73; Conservative 0; Mismatches 64; Indels 0; Gaps 0;

QY  51  CCTGTGCTACTAAGCCAGGTTCTTGTCTTATCTTGTGCTTGGCTATGTTAAAC 110
      |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
Db   257  CCAGTGTGGAGGAGCTGGAGGTCGCTCAAAACTCAGGCAAGATGTATGATGCTTAAC 316

QY  111  CCACCTAACCGTTGTTTGAAGGACACTGATTGTCCAGGTATCAAAAAGTCTGTGAAGGT 170
      |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
Db   317  CCTCCCAATGTCTGCCAGAGGGACGGGCAGTGTGACGGCAAATACAAGTCTGTGAGGTT 376

QY  171  TCCTGCGGTATGGCTTG 187
      |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
Db   377  ATATGTGGGAAAGTCTG 393

RESULT 53
BC028509
LOCUS      BC028509      894 bp      mRNA      linear      ROD 07-AUG-2002
DEFINITION Mus musculus, secretory leukocyte protease inhibitor, clone
            MGC:41142 IMAGE:1513866, mRNA, complete cds.
ACCESSION  BC028509
VERSION     BC028509.1  GI:20306995
KEYWORDS
SOURCE      house mouse.
ORGANISM    Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
REFERENCE   1  (bases 1 to 894)
AUTHORS     Strausberg,R.
TITLE       Direct Submission
JOURNAL     Submitted (23-APR-2002) National Institutes of Health, Mammalian
            Gene Collection (MGC), Cancer Genomics Office, National Cancer
            Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
            USA
REMARK      NIH-MGC Project URL: http://mgc.nci.nih.gov
COMMENT      Contact: MGC help desk
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Marcello Bento Soares, Ph.D.
            cDNA Library Preparation: Soares Laboratory
            cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
            DNA Sequencing by: Genome Sequence Centre,
            BC Cancer Agency, Vancouver, BC, Canada
            info@bcgsc.bc.ca
            Steven Jones, Jennifer Asano, Ian Bosdet, Yaron Butterfield,
            Susanna Chan, Readman Chiu, Chris Fjell, Erin Garland, Ran Guin,
            Letticia Hsiao, Martin Krzywinski, Reta Kutsche, Oliver Lee, Soo
            Sen Lee, Victor Ling, Carrie Mathewson, Candice McLeavy, Steven
            Ness, Pawan Pandoh, Anna-Liisa Prabhu, Parvaneh Saeedi, Jacqueline
            Schein, Duane Smailus, Michael Smith, Lorraine Spence, Jeff Stott,
            Michael Thorne, Miranada Tsai, Natasha van den Bosch, Jill Vardy,
            George Yang, Scott Zuyderduyn, Marco Marra.

Clone distribution: MGC clone distribution information can be found
through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov
Series: IRAC Plate: 67 Row: j Column: 6
This clone was selected for full length sequencing because it
passed the following selection criteria: matched mRNA gi: 6755573.
            Location/Qualifiers
            source       1..894
            /organism="Mus musculus"
            /db_xref="taxon:10090"
            /clone="MGC:41142 IMAGE:1513866"
```


TITLE
JOURNAL
REFERENCE
AUTHORS
TITLE
JOURNAL

Direct Submission

Unpublished
2 (bases 1 to 156460)
Worley,K.C.

Direct Submission

Submitted (15-FEB-2002) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA

REFERENCE
AUTHORS
TITLE
JOURNAL

3 (bases 1 to 156460)
Worley,K.C.

Direct Submission

Submitted (13-JUL-2002) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA

COMMENT

On Jul 12, 2002 this sequence version replaced gi:18767289.

----- Genome Center

Center: Baylor College of Medicine

Center code: BCM

Web site: <http://www.hgsc.bcm.tmc.edu/>

Contact: hgsc-help@bcm.tmc.edu

----- Project Information

Center project name: GRKM

Center clone name: CH230-197B1

----- Summary Statistics

Sequencing vector: Plasmid;

Chemistry: Dye-terminator Big Dye; 100% of reads

Assembly program: Phrap; version 0.990329

Consensus quality: 89134 bases at least Q40

Consensus quality: 94194 bases at least Q30

Consensus quality: 97800 bases at least Q20

* NOTE: Estimated insert size may differ from sequence length
(see http://www.hgsc.bcm.tmc.edu/docs/Genbank_draft_data.html).
* NOTE: This is a 'working draft' sequence. It currently
* consists of 81 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.

1 1169: contig of 1169 bp in length
1170 1269: gap of unknown length
1270 2465: contig of 1196 bp in length
2466 2565: gap of unknown length
2566 3661: contig of 1096 bp in length
3662 3761: gap of unknown length
3762 4956: contig of 1195 bp in length
4957 5056: gap of unknown length
5057 6195: contig of 1139 bp in length
6196 6295: gap of unknown length
6296 7325: contig of 1030 bp in length
7326 7425: gap of unknown length
7426 8480: contig of 1055 bp in length
8481 8580: gap of unknown length
8581 9882: contig of 1302 bp in length
9883 9982: gap of unknown length
9983 11387: contig of 1405 bp in length
11388 11487: gap of unknown length
11488 12675: contig of 1188 bp in length
12676 12775: gap of unknown length
12776 14041: contig of 1266 bp in length
14042 14141: gap of unknown length
14142 15254: contig of 1113 bp in length
15255 15354: gap of unknown length
15355 16409: contig of 1055 bp in length
16410 16509: gap of unknown length
16510 17844: contig of 1335 bp in length
17845 17944: gap of unknown length
17945 19145: contig of 1201 bp in length
19146 19245: gap of unknown length
19246 20466: contig of 1221 bp in length
20467 20566: gap of unknown length

* 20567 22031: contig of 1465 bp in length
* 22032 22131: gap of unknown length
* 22132 23615: contig of 1484 bp in length
* 23616 23715: gap of unknown length
* 23716 25016: contig of 1301 bp in length
* 25017 25116: gap of unknown length
* 25117 26162: contig of 1046 bp in length
* 26163 26262: gap of unknown length
* 26263 27694: contig of 1432 bp in length
* 27695 27794: gap of unknown length
* 27795 29680: contig of 1886 bp in length
* 29681 29780: gap of unknown length
* 29781 30780: contig of 1000 bp in length
* 30781 30880: gap of unknown length
* 30881 31905: contig of 1025 bp in length
* 31906 32005: gap of unknown length
* 32006 33691: contig of 1686 bp in length
* 33692 33791: gap of unknown length
* 33792 35109: contig of 1318 bp in length
* 35110 35209: gap of unknown length
* 35210 36589: contig of 1380 bp in length
* 36590 36689: gap of unknown length
* 36690 38294: contig of 1605 bp in length
* 38295 38394: gap of unknown length
* 38395 40181: contig of 1787 bp in length
* 40182 40281: gap of unknown length
* 40282 41573: contig of 1292 bp in length
* 41574 41673: gap of unknown length
* 41674 43896: contig of 2223 bp in length
* 43897 43996: gap of unknown length
* 43997 45168: contig of 1172 bp in length
* 45169 45268: gap of unknown length
* 45269 47480: contig of 2212 bp in length
* 47481 47580: gap of unknown length
* 47581 48906: contig of 1326 bp in length
* 48907 49006: gap of unknown length
* 49007 50197: contig of 1191 bp in length
* 50198 50297: gap of unknown length
* 50298 51599: contig of 1302 bp in length
* 51600 51699: gap of unknown length
* 51700 52929: contig of 1230 bp in length
* 52930 53029: gap of unknown length
* 53030 55240: contig of 2211 bp in length
* 55241 55340: gap of unknown length
* 55341 56767: contig of 1427 bp in length
* 56768 56867: gap of unknown length
* 56868 58429: contig of 1562 bp in length
* 58430 58529: gap of unknown length
* 58530 59726: contig of 1197 bp in length
* 59727 59826: gap of unknown length
* 59827 61354: contig of 1528 bp in length
* 61355 61454: gap of unknown length
* 61455 63337: contig of 1883 bp in length
* 63338 63437: gap of unknown length
* 63438 64756: contig of 1319 bp in length
* 64757 64856: gap of unknown length
* 64857 66353: contig of 1497 bp in length
* 66354 66453: gap of unknown length
* 66454 67492: contig of 1039 bp in length
* 67493 67592: gap of unknown length
* 67593 69374: contig of 1782 bp in length
* 69375 69474: gap of unknown length
* 69475 71849: contig of 2375 bp in length
* 71850 71949: gap of unknown length
* 71950 73590: contig of 1641 bp in length
* 73591 73690: gap of unknown length
* 73691 75179: contig of 1489 bp in length
* 75180 75279: gap of unknown length
* 75280 78068: contig of 2789 bp in length
* 78069 78168: gap of unknown length
* 78169 80552: contig of 2384 bp in length
* 80553 80652: gap of unknown length
* 80653 82396: contig of 1744 bp in length


```

* 54862 54961: gap of 100 bp
* 54962 55700: contig of 739 bp in length
* 55701 55800: gap of 100 bp

Query Match
Best Local Similarity 16.5%; Score 34; DB 2; Length 61448;
Matches 64; Conservative 0; Mismatches 50; Indels 0; Gaps 0;

QY 81 ATTATCTTGATTGTTGCGCTATGTTAAACCCACCTAACCGTTGTTTGAAGGACACTGAT 140
||||| ||||| ||| ||||| ||||| ||| ||||| ||||| ||||| ||||| |||||
Db 45697 ATTACATTGATGGATTTCCTTATATTAAACCATCCCTGCATTCTCTATGATGAAACCTAT 45638
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 141 TGTCCAGGTATCAAAAAGTGCTGTGAAGGTTCTCTGCGGTATGGCTTGTTCGTT 194
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 45637 TTTGTCAGGATGAATGATTGTTCTTGATGTGTTCTTGATTTCAGTTTGTGAAGTT 45584
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

RESULT 57
AL671891/c
LOCUS
DEFINITION
Mus musculus chromosome X clone RP23-405E21, *** SEQUENCING IN
PROGRESS ***, 9 unordered pieces.
AL671891
AL671891.10 GI:22265400
HTG; HTGS PHASE1; HTGS_ACTIVEFIN; HTGS_DRAFT; HTGS_FULLTOP.
house mouse.
SOURCE
Mus musculus
ORGANISM
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 254836)
Chapman,J.
Direct Submission
Submitted (09-AUG-2002) Wellcome Trust Sanger Institute, Hinxton,
Cambridgeshire, CB10 1SA, UK. E-mail enquiries:
humquery@sanger.ac.uk Clone requests: clonerequest@sanger.ac.uk
On Aug 15, 2002 this sequence version replaced gi:21998201.
----- Genome Center
Center: Wellcome Trust Sanger Institute
Center code: SC
Web site: http://www.sanger.ac.uk
Contact: humquery@sanger.ac.uk
----- Project Information
Center project name: bM405E21
----- Summary Statistics
Assembly program: XGAP4; version 4.5
Chemistry: Dye-terminator; 16% of reads
Chemistry: Dye-terminator Big Dye; 83% of reads
Consensus quality: 250716 bases at least Q40
Consensus quality: 252116 bases at least Q30
Consensus quality: 253330 bases at least Q20
Insert size: 254036; sum-of-contigs
Insert size: 193899; 9.8% error; agarose-fp
Quality coverage: 10.38x in Q20 bases; sum-of-contigs Quality
coverage: 15.09x in Q20 bases; agarose-fp
-----
* NOTE: This is a 'working draft' sequence. It currently
* consists of 9 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
* 1 63987: contig of 63987 bp in length
* 63988 64087: gap of 100 bp
* 64088 195855: contig of 131768 bp in length
* 195856 195955: gap of 100 bp
* 195956 213844: contig of 17889 bp in length
* 213845 213944: gap of 100 bp
* 213945 217297: contig of 3353 bp in length
* 217298 217397: gap of 100 bp
* 217398 220897: contig of 3500 bp in length
* 220898 220997: gap of 100 bp

```

* 220998 223634: contig of 2637 bp in length
* 223635 223734: gap of 100 bp
* 223735 226794: contig of 3060 bp in length
* 226795 226894: gap of 100 bp
* 226895 229736: contig of 2842 bp in length
* 229737 229837: gap of 101 bp
* 229838 254836: contig of 24999 bp in length.

FEATURES

source
1. .254836
/organism="Mus musculus"
/db_xref="taxon:10090"
/chromosome="X"
/clone="RP23-405E21"
/clone_lib="RPCI-23"
1. .63987
/note="assembly_fragment:00034.0"
64088. .195855
/note="assembly_fragment:00034.1"
195956. .213844
/note="assembly_fragment:04368"
213945. .217297
/note="assembly_fragment:05287"
217398. .220897
/note="assembly_fragment:05289"
220998. .223634
/note="assembly_fragment:05296"
223735. .226794
/note="assembly_fragment:05324"
226895. .229736
/note="assembly_fragment:05355"
229837. .254836
/note="assembly_fragment:05381"
BASE COUNT 79852 a 50260 c 48899 g 75017 t 808 others
ORIGIN

Query Match 16.5%; Score 34; DB 2; Length 254836;
Best Local Similarity 56.1%; Pred. No. 7.6;
Matches 64; Conservative 0; Mismatches 50; Indels 0; Gaps 0;
QY 81 ATTATCTTGATTCGTTGCGCTATGTTAAACCCACCTAACCGTGTGTTGAAGGACACTGAT 140
Db 18918 ATTACATTGATGGATTTCCCTTATATTAACCATCCCTGCATTCCTATGAATGAACCTAT 18859
QY 141 TGTCAGGATCAAAAAGTCTGTGAAGGTTCTCGGGTATGGCTTGTTCGTT 194
Db 18858 TTTGTCAGGATGAATGATTGCTCTGATGTGTTCTTGGAATTCAGTTTGGAAGTT 18805

RESULT 58
AL136380
LOCUS AL136380 171656 bp DNA linear PRI 27-FEB-2002
DEFINITION Human DNA sequence from clone RP5-88207 on chromosome 1p33-34.2,
complete sequence.
ACCESSION AL136380
VERSION AL136380.22 GI:19031310
KEYWORDS HTG.
SOURCE human.
ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 171656)

Phillimore,B.
Direct Submission
Submitted (27-FEB-2002) Wellcome Trust Sanger Institute, Hinxton,
Cambridgeshire, CB10 1SA, UK. E-mail enquiries:
humquery@sanger.ac.uk Clone requests: clonerequest@sanger.ac.uk

On Feb 28, 2002 this sequence version replaced gi:18857781.
During sequence assembly data is compared from overlapping clones.
Where differences are found these are annotated as variations
together with a note of the overlapping clone name. Note that the
variation annotation may not be found in the sequence submission
corresponding to the overlapping clone, as we submit sequences with
only a small overlap as described above.

This sequence was finished as follows unless otherwise noted: all
regions were either double-stranded or sequenced with an alternate
chemistry or covered by high quality data (i.e., phred quality >=
30); an attempt was made to resolve all sequencing problems, such
as compressions and repeats; all regions were covered by at least
one plasmid subclone or more than one M13 subclone; and the
assembly was confirmed by restriction digest. The following
abbreviations are used to associate primary accessions given
in the feature table with their source databases: Em:, EMBL; Sw:,
SWISSPROT; Tr:, TREMBL; Wp:, WORMPEP; Information on the WORMPEP
database can be found at
http://www.sanger.ac.uk/Projects/C_elegans/wormpep This sequence
was generated from part of bacterial clone contigs of human
chromosome 1, constructed by the Sanger Centre Chromosome 1 Mapping
Group. Further information can be found at
http://www.sanger.ac.uk/HGP/Chrl
RP5-88207 is from the library RPCI-5 constructed by the group of
Pieter de Jong. For further details see
http://www.chori.org/bacpac/home.htm
VECTOR: pCYPAC2
This sequence is the entire insert of clone RP5-88207 The true left
end of clone RP11-69J16 is at 88191 in this sequence. The true
right end of clone RP11-269F19 is at 13103 in this sequence.

FEATURES

source
1. .171656
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="1"
/map="p33-34.2"
/clone="RP5-88207"
/clone_lib="RPCI-5"
52060
/note="Tandem repeat. Forced join. Gap size estimated to
be approximately 400bp by restriction digest data."

misc_feature

97714. .97772
/note="Sequence from overlapping clone RP11-69J16
(AL359473). Assembly confirmed by restriction digest."

BASE COUNT 48616 a 38050 c 37699 g 47291 t
ORIGIN

Query Match 16.4%; Score 33.8; DB 9; Length 171656;
Best Local Similarity 49.2%; Pred. No. 8.8;
Matches 89; Conservative 0; Mismatches 92; Indels 0; Gaps 0;

QY 1 AATTCGAGCTCGGTACCATACCTGCATATGCTCAAGAACCAAGTTAAAGGTCCTGTGTCTA 60
Db 133686 AAATCAATGGGCTTGGATTGGACATCTCTCAATAAAAGGGTTAAAGGCTGATGGACC 133745

QY 61 CTAAGCCAGGTTCTTGTCCCTATTATCTTGATTGCTGGCTATGTTAAACCCACCTAAC 120
Db 133746 TAAAGCCGTGACTTGAATTTGATCAAGATACGCTGCCTTAAGTTCTCTTCATTACACA 133805

QY 121 GTTGTTTGAAGGACACTGATTGTCCAGGTATCAAAAAGTGTGAAGTTCTCTGGGTA 180
Db 133806 GATGATCCTAGGTAATTGATAGATCCTGTGGTTCAACTTGATTCTAGATAGAAGCTGGA 133865

QY 181 T 181

Db 133866 T 133866

RESULT 59

AC116110/c

LOCUS AC116110

DEFINITION Mus musculus clone RP23-19B12, WORKING DRAFT SEQUENCE, 12 ordered
pieces.

ACCESSION AC116110

VERSION AC116110.2 GI:22474982

KEYWORDS HTG; HTGS_PHASE2; HTGS_DRAFT; HTGS_FULLTOP.

SOURCE house mouse.

ORGANISM Mus musculus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

AC116110 178652 bp DNA linear HTG 24-AUG-2002
Mus musculus clone RP23-19B12, WORKING DRAFT SEQUENCE, 12 ordered
pieces.

REFERENCE
AUTHORS
JOURNAL
REFERENCE
AUTHORS

1 (bases 1 to 178652)
Birren,B., Nusbaum,C. and Lander,E.
Mus musculus, clone RP23-19B12
Unpublished

2 (bases 1 to 178652)
Birren,B., Linton,L., Nusbaum,C., Lander,E., Ali,A., Allen,N.,
Anderson,S., Barna,N., Bastien,V., Bloom,T., Boguslavskiy,L.,
Boukhgalter,B., Brown,A., Camarata,J., Campopiano,A., Chang,J.,
Chazaro,B., Choepel,Y., Colangelo,M., Collins,S., Collamore,A.,
Cook,A., Cooke,P., DeArellano,K., Dewar,K., Diaz,J.S., Dodge,S.,
Faro,S., Ferreira,P., FitzHugh,W., Gage,D., Galagan,J., Gardyna,S.,
Ginde,S., Gord,S., Goyette,M., Graham,L., Grand-Pierre,N.,
Hagos,B., Horton,L., Hulme,W., Iliev,I., Johnson,R., Jones,C.,
Kamat,A., Karatas,A., Kells,C., LaRocque,K., Lamazares,R.,
Landers,T., Lehoczy,J., Levine,R., Lindblad-Toh,K., Liu,G.,
MacLean,C., Macdonald,P., Major,J., Marquis,N., Matthews,C.,
McCarthy,M., McEwan,P., McKernan,K., Meldrim,J., Meneus,L.,
Mihova,T., Mlenga,V., Murphy,T., Naylor,T., Nguyen,C., Nicol,R.,
Norbu,C., Norman,C.H., O'Connor,T., O'Donnell,P., O'Neil,D.,
Oliver,J., Peterson,K., Phunkhang,P., Pierre,N., Pollara,V.,
Raymond,C., Retta,R., Rieback,M., Riley,R., Rise,C., Rogov,P.,
Roman,J., Rosetti,M., Roy,A., Santos,R., Schauer,S., Schupback,R.,
Seaman,S., Severy,P., Spencer,B., Stange-Thomann,N., Stojanovic,N.,
Strauss,N., Subramanian,A., Talamas,J., Tesfaye,S., Theodore,J.,
Topham,K., Travers,M., Travis,N., Trigilio,J., Vassiliev,H.,
Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Ye,W.J., Young,G.,
Zainoun,J., Zembek,L., Zimmer,A. and Zody,M.

TITLE
JOURNAL
REFERENCE
AUTHORS

Submitted (25-MAR-2002) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA

3 (bases 1 to 178652)
Birren,B., Nusbaum,C., Lander,E., Ali,A., Allen,N., Anderson,S.,
Barna,N., Bastien,V., Bloom,T., Boguslavskiy,L., Boukhgalter,B.,
Camarata,J., Chang,J., Chazaro,B., Choepel,Y., Collamore,A.,
Cook,A., Cooke,P., DeArellano,K., Dewar,K., Diaz,J.S., Dodge,S.,
Faro,S., Ferreira,P., FitzGerald,M., Gage,D., Galagan,J.,
Gardyna,S., Gord,S., Graham,L., Grand-Pierre,N., Hagos,B.,
Horton,L., Hulme,W., Iliev,I., Johnson,R., Jones,C., Kamat,A.,
Karatas,A., Kells,C., Landers,T., Levine,R., Lindblad-Toh,K.,
Liu,G., MacLean,C., Macdonald,P., Major,J., Matthews,C.,
McCarthy,M., Meldrim,J., Meneus,L., Mihova,T., Mlenga,V.,
Murphy,T., Naylor,J., Nguyen,C., Nicol,R., Norbu,C., Norman,C.H.,
O'Connor,T., O'Donnell,P., O'Neil,D., Oliver,J., Peterson,K.,
Phunkhang,P., Pierre,N., Raymond,C., Retta,R., Rise,C., Rogov,P.,
Roman,J., Roy,A., Schauer,S., Schupback,R., Seaman,S., Severy,P.,
Smith,C., Spencer,B., Stange-Thomann,N., Stojanovic,N., Talamas,J.,
Tesfaye,S., Theodore,J., Topham,K., Travers,M., Vassiliev,H.,
Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Young,G., Zainoun,J.,
Zembek,L., Zimmer,A. and Zody,M.

TITLE
JOURNAL
COMMENT

Direct Submission
Submitted (24-AUG-2002) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA
On Aug 24, 2002 this sequence version replaced gi:19703221.
All repeats were identified using RepeatMasker:
Smit, A.F.A. & Green, P. (1996-1997)
<http://ftp.genome.washington.edu/RM/RepeatMasker.html>

----- Genome Center
Center: Whitehead Institute/ MIT Center for Genome Research
Center code: WIBR
Web site: <http://www-seq.wi.mit.edu>
Contact: sequence_submissions@genome.wi.mit.edu
----- Project Information
Center project name: l22837
Center clone name: 19_B_12
----- Summary Statistics

Sequencing vector: Plasmid; n/a; 100% of reads
Chemistry: Dye-terminator Big Dye; 100% of reads
Assembly program: Phrap; version 0.960731
Consensus quality: 175241 bases at least Q40
Consensus quality: 176762 bases at least Q30
Consensus quality: 177264 bases at least Q20
Insert size: 176000; agarose-fp
Insert size: 177552; sum-of-contigs

Quality coverage: 8.4 in Q20 bases; agarose-fp
Quality coverage: 8.3 in Q20 bases; sum-of-contigs

* NOTE: This is a 'working draft' sequence. It currently
* consists of 12 contigs. Gaps between the contigs
* are represented as runs of N. The order of the pieces
* is believed to be correct as given, however the sizes
* of the gaps between them are based on estimates that have
* provided by the submitter.

* This sequence will be replaced
* by the finished sequence as soon as it is available and
* the accession number will be preserved.

* 1 816: contig of 816 bp in length
* 817 916: gap of 100 bp
* 917 1985: contig of 1069 bp in length
* 1986 2085: gap of 100 bp
* 2086 2896: contig of 811 bp in length
* 2897 2996: gap of 100 bp
* 2997 7223: contig of 4227 bp in length
* 7224 7323: gap of 100 bp
* 7324 13183: contig of 5860 bp in length
* 13184 13283: gap of 100 bp
* 13284 126261: contig of 112978 bp in length
* 126262 126361: gap of 100 bp
* 126362 133596: contig of 7235 bp in length
* 133597 133696: gap of 100 bp
* 133697 139614: contig of 5918 bp in length
* 139615 139714: gap of 100 bp
* 139715 145185: contig of 5471 bp in length
* 145186 145285: gap of 100 bp
* 145286 154037: contig of 8752 bp in length
* 154038 154137: gap of 100 bp
* 154138 175237: contig of 21100 bp in length
* 175238 175337: gap of 100 bp
* 175338 178652: contig of 3315 bp in length.

FEATURES
source

1. 178652
/organism="Mus musculus"
/db xref="taxon:10090"
/clone="RP23-19B12"
/clone_lib="RPCI-23 Female Mouse BAC"

misc_feature

1. 816
/note="assembly_fragment"
clone_end:SP6
vector_side:left"

misc_feature

917. 1985
/note="assembly_fragment"
2086. 2896
/note="assembly_fragment"

misc_feature

2997. 7223
/note="assembly_fragment"
7324. 13183
/note="assembly_fragment"

misc_feature

13284. 126261
/note="assembly_fragment"
126362. 133596
/note="assembly_fragment"

misc_feature

133697. 139614
/note="assembly_fragment"
139715. 145185
/note="assembly_fragment"

misc_feature

145286. 154037
/note="assembly_fragment"
154138. 175237
/note="assembly_fragment"

misc_feature

175338. 178652
/note="assembly_fragment"
clone_end:T7
vector_side:right"

BASE COUNT 53824 a 33935 c 32001 g 57792 t 1100 others
ORIGIN

Query Match

16.4%; Score 33.8; DB 2; Length 178652;

Best Local Similarity 55.6%; Pred. No. 8.8; Matches 65; Conservative 0; Mismatches 52; Indels 0; Gaps 0;

QY 88 TGATTCGTTGGCTATGTTAAACCCACCTAACCGTTGTTTGAAGGACACTGATTGTCCAG 147
|||||
Db 102074 TGATTCTCTGTTATACCTATCAACTTGTGAACCTCTAACATCTTGACAGTGAAAGCCAAAG 102015
|||||

QY 148 GTATCAAAAAGTCTGTGAAGGTTCTCTGCGGTATGGCTTGTTCGTTCCACAATAAT 204
|||||
Db 102014 CTTTCTCCCATCTCTGTGAAAAGTCTTGGAGGCAAGGTTATGATGTTCCACTATACT 101958
|||||

RESULT 60
AF002719 AF002719 2435 bp DNA linear ROD 26-JAN-1999
LOCUS Mus musculus secretory leukoprotease inhibitor gene, complete cds.
DEFINITION AF002719
ACCESSION AF002719
VERSION AF002719.1 GI:4100898
KEYWORDS
SOURCE Mus musculus.
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 2435)
AUTHORS Kikuchi,T., Abe,T., Hoshi,S., Matsubara,N., Tominaga,Y., Satoh,K.
and Nukiwa,T.
TITLE Structure of the murine secretory leukoprotease inhibitor (Slpi)
gene and chromosomal localization of the human and murine SLPI
genes
JOURNAL Am. J. Respir. Cell Mol. Biol. 19 (6), 875-880 (1998)
MEDLINE 99061820
PUBMED 9843921
REFERENCE 2 (bases 1 to 2435)
AUTHORS Kikuchi,T., Abe,T., Hoshi,S., Matsubara,N., Tominaga,Y., Satoh,K.
and Nukiwa,T.
TITLE Direct Submission
JOURNAL Submitted (06-MAY-1997) Department of Respiratory Oncology and
Molecular Medicine, Institute of Development, Aging and Cancer,
Tohoku University, 4-1 Seiryomachi, Aoba-ku, Sendai, Miyagi-ken
980-77, Japan
FEATURES
source Location/Qualifiers
1..2435
/organism="Mus musculus"
/strain="129SV"
/db_xref="taxon:10090"
/chromosome="2"
join(101..205,781..942,1368..1517,2080..2322)
/product="secretory leukoprotease inhibitor"
join(121..205,781..942,1368..1516)
/codon_start=1
/product="secretory leukoprotease inhibitor"
/protein_id="AAD09307.1"
/db_xref="GI:4100899"
/translation="MKSCGLLPFTVLALGILAPWTVEGGKNDKIGACPAKKPAQC
LKLEKPCRTDWECPGKQRCQDAGCKVCNVPVPIKPVWRKPGRCVKQTQARCMMLNP
PNVCQRDGGQCDGKYKCEGICGKVCCLPPM"
BASE COUNT 509 a 600 c 652 g 674 t
ORIGIN
Query Match 16.3%; Score 33.6; DB 10; Length 2435;
Best Local Similarity 52.9%; Pred. No. 11;
Matches 72; Conservative 0; Mismatches 64; Indels 0; Gaps 0;

QY 52 CTGTGCTACTAAGCCAGGTTCTTGTCCTATTATCTTGATGCTGTTGGCTATGTTAAACC 111
|||||
Db 1365 CAGTGTGGAGGAAGCCTGGGAGGTGCTCAAAACTCAGGCAAGATGATGATGCTTAACC 1424
|||||

QY 112 CACCTAACCGTTGTTGAAGGACACTGATGTTCCAGGTATCAAAAAGTCTGTGAAGGTT 171
|||||
Db 1425 CTCCCAATGTCTGCCAGAGGACGGGAGGTGTGACGGCAATAACAAGTCTGTGAGGTA 1484
|||||

QY 172 CCTGCGGTATGGCTTG 187
|||||

Db 1485 TATGTGGGAAAGTCTG 1500

RESULT 61
AL662897 84987 bp DNA linear PRI 22-JAN-2002
LOCUS Human DNA sequence from clone RP11-309C8 on chromosome X, complete
DEFINITION sequence.
ACCESSION AL662897
VERSION AL662897.6 GI:18307360
KEYWORDS HTG.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 84987)
AUTHORS Howden,P.
TITLE Direct Submission
JOURNAL Submitted (22-JAN-2002) Wellcome Trust Sanger Institute, Hinxton,
Cambridgeshire, CB10 1SA, UK. E-mail enquiries:
humquerry@sanger.ac.uk Clone requests: clonerequest@sanger.ac.uk
On Jan 23, 2002 this sequence version replaced gi:18250572.
During sequence assembly data is compared from overlapping clones.
Where differences are found these are annotated as variations
together with a note of the overlapping clone name. Note that the
variation annotation may not be found in the sequence submission
corresponding to the overlapping clone, as we submit sequences with
only a small overlap as described above.
This sequence was finished as follows unless otherwise noted: all
regions were either double-stranded or sequenced with an alternate
chemistry or covered by high quality data (i.e., phred quality >=
30); an attempt was made to resolve all sequencing problems, such
as compressions and repeats; all regions were covered by at least
one plasmid subclone or more than one M13 subclone; and the
assembly was confirmed by restriction digest. The following
abbreviations are used to associate primary accession numbers given
in the feature table with their source databases: Em., EMBL; Sw.,
SWISSPROT; Tr., TrEMBL; Wp., WORMPEP; Information on the WORMPEP
database can be found at
http://www.sanger.ac.uk/Projects/C_elegans/wormpep This sequence
was generated from part of bacterial clone contigs of human
chromosome X, constructed by the Sanger Centre Chromosome X Mapping
Group. Further information can be found at
http://www.sanger.ac.uk/HGP/ChrX
RP11-309C8 is from the library RPCI-11.2 constructed by the group
of Pieter de Jong. For further details see
http://www.chori.org/bacpac/home.htm
VECTOR: pBACE3.6
IMPORTANT: This sequence is not the entire insert of clone
RP11-309C8 It may be shorter because we sequence overlapping
sections only once, except for a short overlap.
The true left end of clone RP1-198G23 is at 82988 in this sequence.
The true right end of clone RP11-378H3 is at 2000 in this sequence.
FEATURES
source Location/Qualifiers
1..84987
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="X"
/clone="RP11-309C8"
/clone_lib="RPCI-11.2"
BASE COUNT 26631 a 15105 c 15585 g 27666 t
ORIGIN
Query Match 16.3%; Score 33.6; DB 9; Length 84987;
Best Local Similarity 57.7%; Pred. No. 10;
Matches 60; Conservative 0; Mismatches 44; Indels 0; Gaps 0;

QY 53 TGTGCTACTAAGCCAGGTTCTTGTCCTATTATCTTGATTCGTTGCGTATGTTAAACC 112
|||||
Db 49793 TCTGCTCTTTATCTATTTATGCGGTATATTATGTTGATTGTTTGCATGTTAAACCA 49852
|||||

QY 113 ACCTAACCGTTGTTGAAGGACACTGATGTTCCAGGTATCAAAA 156
|||||

corresponding to the overlapping clone, as we submit sequences with only a small overlap as described above.

The following abbreviations are used to associate primary accession numbers given in the feature table with their source databases:

Em:, EMBL; Sw:, SWISSPROT; Tr:, TREMBL; Wp:, WORMPEP; Information on the WORMPEP database can be found at

http://www.sanger.ac.uk/Projects/C_elegans/wormpep RP1-319F24 is

from the library RPCI-1 constructed at the Roswell Park Cancer

Institute by the group of Pieter de Jong. For further details see

<http://bacpac.med.buffalo.edu/>

VECTOR: pCYPAC2

IMPORTANT: This sequence is not the entire insert of clone

RP1-319F24 It may be shorter because we only sequence overlapping

sections once, or longer because we arrange for a small overlap

between neighbouring submissions.

The true left end of clone RP3-508I15 is at 28416 in this sequence.

The true right end of clone RP1-199H16 is at 100 in this sequence.

FEATURES

source

```
1. .28515
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="22"
/map="q12.3-13.31"
/clone="RP1-319F24"
/clone_lib="RPCI-1"
2. .157
/note="AluSx repeat: matches 1. .154 of consensus"
1186. .1231
/note="23 copies 2 mer ta 80 conserved"
1233. .1534
/note="AluSx repeat: matches 1. .297 of consensus"
complement(1558. .1698)
/note="Single clone region"
1573. .1867
/note="AluSx repeat: matches 1. .302 of consensus"
1926. .2825
/note="L2 repeat: matches 1747. .2746 of consensus"
3082. .3285
/note="MER20 repeat: matches 2. .218 of consensus"
3763. .4070
/note="AluSx repeat: matches 1. .308 of consensus"
4108. .4401
/note="AluSx repeat: matches 1. .299 of consensus"
4944. .5249
/note="AluSx repeat: matches 3. .310 of consensus"
5469. .5542
/note="MIR repeat: matches 8. .81 of consensus"
5543. .5849
/note="AluSx repeat: matches 1. .308 of consensus"
5890. .6084
/note="MIR repeat: matches 68. .262 of consensus"
6281. .6486
/note="AluSx repeat: matches 1. .206 of consensus"
6487. .6551
/note="L1 repeat: matches 2983. .3047 of consensus"
6547. .6610
/note="L1M4 repeat: matches 2885. .2950 of consensus"
6863. .6908
/note="MER74B repeat: matches 1. .46 of consensus"
6952. .7069
/note="MER74B repeat: matches 153. .271 of consensus"
7070. .7342
/note="AluJo repeat: matches 24. .296 of consensus"
7343. .7499
/note="MER74B repeat: matches 271. .442 of consensus"
7794. .8096
/note="AluJo repeat: matches 1. .305 of consensus"
8386. .8421
/note="18 copies 2 mer ag 88 conserved"
8570. .8684
/note="MIR repeat: matches 21. .132 of consensus"
8814. .8912
/note="L2 repeat: matches 2540. .2678 of consensus"
```

```
repeat_region
8944. .9016
/note="MLT1J repeat: matches 430. .510 of consensus"
9017. .9241
/note="AluJb repeat: matches 84. .300 of consensus"
9245. .9331
/note="MLT1J repeat: matches 350. .440 of consensus"
9332. .9625
/note="AluSx repeat: matches 3. .296 of consensus"
9626. .9683
/note="MLT1J repeat: matches 290. .350 of consensus"
9684. .9989
/note="AluSx repeat: matches 1. .310 of consensus"
9990. .10296
/note="MLT1J repeat: matches 11. .290 of consensus"
10305. .10441
/note="L2 repeat: matches 2418. .2548 of consensus"
10656. .10969
/note="AluSg repeat: matches 1. .313 of consensus"
10971. .11220
/note="AluJb repeat: matches 2. .300 of consensus"
11393. .11517
/note="MIR repeat: matches 79. .207 of consensus"
11525. .11617
/note="AluSp/q repeat: matches 202. .308 of consensus"
11678. .11974
/note="AluSg repeat: matches 1. .298 of consensus"
12666. .13415
/note="L2 repeat: matches 1678. .2473 of consensus"
complement(14319. .14778)
/note="match: GSS: Em:AQ164414"
14607. .14767
/note="MIR repeat: matches 90. .260 of consensus"
14776. .14976
/note="AluSg/x repeat: matches 85. .291 of consensus"
14779. .15152
/note="match: GSS: Em:AQ122011"
14978. .15035
/note="MIR repeat: matches 7. .68 of consensus"
15807. .15876
/note="MIR repeat: matches 185. .256 of consensus"
15870. .15956
/note="L2 repeat: matches 2667. .2748 of consensus"
15957. .16218
/note="AluSx repeat: matches 3. .265 of consensus"
16219. .16794
/note="L2 repeat: matches 2021. .2667 of consensus"
16812. .17288
/note="L2 repeat: matches 1371. .1777 of consensus"
17289. .17584
/note="AluSg repeat: matches 12. .304 of consensus"
17585. .18113
/note="L2 repeat: matches 623. .1371 of consensus"
18076. .18162
/note="MIR repeat: matches 81. .164 of consensus"
18190. .18491
/note="AluJ repeat: matches 1. .302 of consensus"
18616. .18869
/note="AluJo repeat: matches 1. .308 of consensus"
19009. .19579
/note="MER34 repeat: matches 5. .543 of consensus"
19628. .19839
/note="MER58A repeat: matches 1. .224 of consensus"
19981. .20210
/note="L2 repeat: matches 2198. .2414 of consensus"
20211. .20520
/note="AluSx repeat: matches 1. .310 of consensus"
20521. .20795
/note="L2 repeat: matches 1881. .2198 of consensus"
21420. .21726
/note="AluJ repeat: matches 1. .302 of consensus"
21952. .22274
/note="AluSg repeat: matches 1. .308 of consensus"
22393. .22509
```


100

Db 81022 TTTTCACTAGACTCAAGTCCAGCTCCAATCTCTCTCCACCAGAGTCTGGGCAATTCTC 81081

QY 87 TTGATTCTGTTGCGCTATGTTAAACCCACCTAAC 120

Db 81082 TTGATTCACTCTCTGACAAAAACAACAAC 81115

RESULT 69

AF375596S2 17515 bp DNA linear ROD 20-MAY-2001

LOCUS Mus musculus medium and short chain L-3-hydroxyacyl-Coenzyme A

DEFINITION dehydrogenase (Mschad) gene, exons 2 through 8, and complete cds;

ACCESSION AF375597

VERSION AF375597.1 GI:14150816

KEYWORDS nuclear gene for mitochondrial product.

SEGMENT 2 of 2

SOURCE Mus musculus.

ORGANISM Mus musculus

REFERENCE 1 (bases 1 to 17515)

AUTHORS O'Brien, L.K., Sims, H.F. and Strauss, A.W.

TITLE Mouse medium and short chain L-3-hydroxyacyl-Coenzyme A dehydrogenase gene

JOURNAL Unpublished

REFERENCE 2 (bases 1 to 17515)

AUTHORS O'Brien, L.K., Sims, H.F. and Strauss, A.W.

TITLE Direct Submission

JOURNAL Submitted (02-MAY-2001) Pediatrics, Vanderbilt Medical Center, 1161 21st Avenue South, D-2220 MCN, Nashville, TN 37232, USA

FEATURES

source 1. .17515

gene /organism="Mus musculus"

order(AF375596.1:6458. .10094,1. .17515)

join(AF375596.1:6458. .6609,1329. .1457,2698. .2855, 5986. .6112,8631. .8720,10231. .10303,15582. .15698, 16954. .17515)

gene="Mschad"

product="medium and short chain L-3-hydroxyacyl-Coenzyme A dehydrogenase"

join(AF375596.1:6478. .6609,1329. .1457,2698. .2855, 5986. .6112,8631. .8720,10231. .10303,15582. .15698, 16954. .17072)

gene="Mschad"

note="L-3-hydroxyacyl-Coenzyme A dehydrogenase; NAD-dependent dehydrogenase; specific for medium and short chain fatty acids; SCHAD; HADH"

codon_start=1

product="medium and short chain L-3-hydroxyacyl-Coenzyme A dehydrogenase"

protein_id="AAK54642.1"

db_xref="GI:14150817"

translation="MAFVTRQFLRSMSSSSASAAKKILIKHVTVIGGLMGAGIAQ VAAATGHTVVLVDQTEDILAKSKKGIIEESLKRMAKKFTENPKAGDEFVEKTLSCLS STDAASVHSTDLVVEAIVENLKLKNEFLQRLDKFAAEHTIPASNTSSLOITNINAT TRQDRFAGLHFPNPVPMKLVVEIKTPMTSQTFESLVDFCKTLGHPVSKDTPGFI VNRLLVPYLFAVRLHERGDASKEDIDTAMKLGAGYPMGPPELLDYVGLDITTKFLDGM WHMEPENPLFQPSPSMNNLVAQKLGKKTGEGFYKYK"

exon 1329. .1427

gene="Mschad"

number=2

exon 2698. .2855

gene="Mschad"

number=3

exon 5986. .6112

gene="Mschad"

number=4

exon 8631. .8720

gene="Mschad"

exon /number=5

10231. .10303

gene="Mschad"

number=6

15582. .15698

gene="Mschad"

number=7

16954. .17515

gene="Mschad"

number=8

BASE COUNT 3910 a 4109 c 4462 g 5026 t 8 others

ORIGIN

Query Match 16.0%; Score 33; DB 10; Length 17515;

Best Local Similarity 52.6%; Pred. No. 16;

Matches 72; Conservative 0; Mismatches 65; Indels 0; Gaps 0;

QY 69 GGTTCCTGTCCTATTATCTTGATTGCTTGGCGCTATGTTAAACCCACCTAACCGTTGTTG 128

Db 9488 GTTTGTTTGTGTTTGTGTTTGTGTTTGTGTTTGTGTTTGTGTTTGTGTTTGTGTTG 9547

QY 129 AAGGACACTGATTGTCCAGGTATCAAAAAGTCTGTAAGGTTCTCTCGGATGCTTGT 188

Db 9548 TCCTGAACCCATGAGTCAGCTCTCTTTAGCTACTGGGCCCTCTACTGCCTAATGGCTTGT 9607

QY 189 TTCGTTCCACAATAATA 205

Db 9608 TTCAGAGAACAAACTA 9624

RESULT 70

AC119403

LOCUS Homo sapiens chromosome 19 clone LLNLR-265A7, complete sequence.

DEFINITION AC119403

ACCESSION AC119403

VERSION AC119403.2 GI:21622702

KEYWORDS HTG.

SOURCE human.

ORGANISM Homo sapiens

REFERENCE 1 (bases 1 to 40223)

AUTHORS DOE Joint Genome Institute and Stanford Human Genome Center.

TITLE Direct Submission

JOURNAL Unpublished

REFERENCE 2 (bases 1 to 40223)

AUTHORS DOE Joint Genome Institute.

TITLE Direct Submission

JOURNAL Submitted (26-APR-2002) Production Sequencing Facility, DOE Joint Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA

REFERENCE 3 (bases 1 to 40223)

AUTHORS DOE Joint Genome Institute and Stanford Human Genome Center.

TITLE Direct Submission

JOURNAL Submitted (28-JUN-2002) DOE Joint Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA

COMMENT On Jun 28, 2002 this sequence version replaced gi:20330817. Draft Sequence Produced by DOE Joint Genome Institute www.jgi.doe.gov

Finishing Completed at Stanford Human Genome Center

www-shgc.stanford.edu

Quality: Phrap Quality >=40 99.6% of Sequence;

Estimated Total Number of Errors is 0.1.

FEATURES

source 1. .40223

organism="Homo sapiens"

db_xref="taxon:9606"

chromosome="19"

clone="LLNLR-265A7"

BASE COUNT 10097 a 9482 c 9319 g 11325 t

ORIGIN

Query Match 16.0%; Score 33; DB 9; Length 40223;

Best Local Similarity 54.5%; Pred. No. 16;

COMMENT

On May 4, 2002 this sequence version replaced gi:20389394.
All repeats were identified using RepeatMasker:
Smit, A.F.A. & Green, P. (1996-1997)
<http://ftp.genome.washington.edu/RM/RepeatMasker.html>
----- Genome Center
Center: Whitehead Institute/ MIT Center for Genome Research
Center code: WIBR
Web site: <http://www-seq.wi.mit.edu>
Contact: sequence_submissions@genome.wi.mit.edu
----- Project Information
Center project name: L19584
Center clone name: 157 H_14

----- Summary Statistics
Sequencing vector: Plasmid; n/a; 100% of reads
Chemistry: Dye-terminator Big Dye; 100% of reads
Assembly program: Phrap; version 0.960731
Consensus quality: 200834 bases at least Q40
Consensus quality: 204678 bases at least Q30
Consensus quality: 206239 bases at least Q20
Insert size: 188000; agarose-fp
Insert size: 209203; sum-of-contigs
Quality coverage: 8.3 in Q20 bases; agarose-fp
Quality coverage: 7.5 in Q20 bases; sum-of-contigs

* NOTE: This is a 'working draft' sequence. It currently
* consists of 46 contigs. Gaps between the contigs
* are represented as runs of N. The order of the pieces
* is believed to be correct as given, however the sizes
* of the gaps between them are based on estimates that have
* provided by the submitter.

* This sequence will be replaced
* by the finished sequence as soon as it is available and
* the accession number will be preserved.

1 651: contig of 651 bp in length
652 751: gap of 100 bp
752 1458: contig of 707 bp in length
1459 1558: gap of 100 bp
1559 2267: contig of 709 bp in length
2268 2367: gap of 100 bp
2368 3028: contig of 661 bp in length
3029 3128: gap of 100 bp
3129 3802: contig of 674 bp in length
3803 3902: gap of 100 bp
3903 4563: contig of 661 bp in length
4564 4663: gap of 100 bp
4664 5327: contig of 664 bp in length
5328 5427: gap of 100 bp
5428 6088: contig of 661 bp in length
6089 6188: gap of 100 bp
6189 6866: contig of 678 bp in length
6867 6966: gap of 100 bp
6967 7647: contig of 681 bp in length
7648 7747: gap of 100 bp
7748 8439: contig of 692 bp in length
8440 8539: gap of 100 bp
8540 9198: contig of 659 bp in length
9199 9298: gap of 100 bp
9299 9711: contig of 413 bp in length
9712 9811: gap of 100 bp
9812 10478: contig of 667 bp in length
10479 10578: gap of 100 bp
10579 11249: contig of 671 bp in length
11250 11349: gap of 100 bp
11350 11996: contig of 647 bp in length
11997 12096: gap of 100 bp
12097 12785: contig of 689 bp in length
12786 12885: gap of 100 bp
12886 13536: contig of 651 bp in length
13537 13636: gap of 100 bp
13637 14292: contig of 656 bp in length
14293 14392: gap of 100 bp
14393 15085: contig of 693 bp in length
15086 15185: gap of 100 bp

* 15186 15835: contig of 650 bp in length
* 15836 15935: gap of 100 bp
* 15936 17225: contig of 1290 bp in length
* 17226 17325: gap of 100 bp
* 17326 19181: contig of 1856 bp in length
* 19182 19281: gap of 100 bp
* 19282 21156: contig of 1875 bp in length
* 21157 21256: gap of 100 bp
* 21257 24719: contig of 3463 bp in length
* 24720 24819: gap of 100 bp
* 24820 27090: contig of 2271 bp in length
* 27091 27190: gap of 100 bp
* 27191 30340: contig of 3150 bp in length
* 30341 30440: gap of 100 bp
* 30441 32803: contig of 2363 bp in length
* 32804 32903: gap of 100 bp
* 32904 35837: contig of 2934 bp in length
* 35838 35937: gap of 100 bp
* 35938 39727: contig of 3790 bp in length
* 39728 39827: gap of 100 bp
* 39828 43907: contig of 4080 bp in length
* 43908 44007: gap of 100 bp
* 44008 49229: contig of 5222 bp in length
* 49230 49329: gap of 100 bp
* 49330 52432: contig of 3103 bp in length
* 52433 52532: gap of 100 bp
* 52533 71268: contig of 18736 bp in length
* 71269 71368: gap of 100 bp
* 71369 78519: contig of 7151 bp in length
* 78520 78619: gap of 100 bp
* 78620 87207: contig of 8588 bp in length
* 87208 87307: gap of 100 bp
* 87308 96949: contig of 9642 bp in length
* 96950 97049: gap of 100 bp
* 97050 104314: contig of 7265 bp in length
* 104315 104414: gap of 100 bp
* 104415 113053: contig of 8639 bp in length
* 113054 113153: gap of 100 bp
* 113154 125823: contig of 12670 bp in length
* 125824 125923: gap of 100 bp
* 125924 139598: contig of 13675 bp in length
* 139599 139698: gap of 100 bp
* 139699 153566: contig of 13868 bp in length
* 153567 153666: gap of 100 bp
* 153667 167089: contig of 13423 bp in length
* 167090 167189: gap of 100 bp
* 167190 184611: contig of 17422 bp in length
* 184612 184711: gap of 100 bp
* 184712 205546: contig of 20835 bp in length
* 205547 205646: gap of 100 bp
* 205647 213703: contig of 8057 bp in length.

FEATURES

source

1. .213703
/organism="Mus musculus"
/db_xref="taxon:10090"
/clone="RP23-157H14"
/clone lib="RPC1-23 Female Mouse BAC"

misc_feature

1. .651

misc_feature

752. .1458

Query Match 16.0%; Score 33; DB 2; Length 213703;
Best Local Similarity 50.3%; Pred. No. 16;
Matches 81; Conservative 0; Mismatches 80; Indels 0; Gaps 0;

QY 4 TCGAGCTCGGTACCATACCTGCATATGCTCAAGAACAGTTAAAGGTCCTGTCTACTA 63

Db 116461 TCTATCAGGTACTATTTCACAATAATCAACAAACAGACTTTGGTTGTACTGTA 116520

QY 64 AGCCAGGTTCTTGTCTATTATCTTGATTCGTTGGCTATGTTAAACCCACCGTT 123

Db 116521 CAGCTGATTGTTTCACTACTTGGATGTAAGGTGTCATATTGTTGACACAAATATTGTA 116580

repeat_region	/rpt_family="LTR37A"	5845. .6070	Query Match	15.9%;	Score 32.8;	DB 9;	Length 143620;
repeat_region	/rpt_family="MER92C"	6093. .6156	Best Local Similarity	54.0%;	Pred. No. 18;		
repeat_region	/rpt_family="MER92C"	6706. .6751	Matches	67;	Conservative	0;	Mismatches 57; Indels 0; Gaps 0;
repeat_region	/rpt_family="AT_rich"	complement(7098. .7306)	51	CCTGTGCTACTAAGCCAGGTTCTTGCTCTATTATCTTGATTCGTTGCGCTATGTTAAAC	110		
repeat_region	/rpt_family="MIR"	7512. .8234	Db 137996	CTTTTGCTTTTATTCCATTATATGAAGTATTACATTGATTGTTTTGTTATGTTAAAC	137937		
repeat_region	/rpt_family="LTR41"	complement(8687. .8831)	Qy	111	CCACCTAACCGTTGTTTGAAGGACACTGATTGTCCAGGTATCAAAAAGTCTGTGAAGGT	170	
repeat_region	/rpt_family="MIR"	8861. .8962					
repeat_region	/rpt_family="L2"	9468. .9913					
repeat_region	/rpt_family="L1MEC"	10108. .10685					
repeat_region	/rpt_family="L1MEC"	10787. .11269					
repeat_region	/rpt_family="L2"	11878. .11937					
repeat_region	/rpt_family="GA-rich"	complement(12456. .12699)					
repeat_region	/rpt_family="AluSc"	complement(12811. .13551)					
repeat_region	/rpt_family="LTR48"	13989. .14009					
repeat_region	/rpt_family="AT_rich"	complement(15576. .15766)					
repeat_region	/rpt_family="MER5A"	complement(15789. .15941)					
repeat_region	/rpt_family="MIR"	complement(15971. .16065)					
repeat_region	/rpt_family="MIR"	16503. .16920					
repeat_region	/rpt_family="MSTA"	16923. .17892					
repeat_region	/rpt_family="MSTA-internal"	complement(17893. .18655)					
repeat_region	/rpt_family="LTR17"	18656. .19271					
repeat_region	/rpt_family="MSTA-internal"	19272. .19640					
repeat_region	/rpt_family="MSTA"	19675. .19738					
repeat_region	/rpt_family="MIR"	19823. .19843					
repeat_region	/rpt_family="AT_rich"	20062. .20084					
repeat_region	/rpt_family="(CAAAA)n"	complement(21139. .21789)					
repeat_region	/rpt_family="L1MC3"	22734. .22781					
repeat_region	/rpt_family="MIR"	22739. .22791					
repeat_region	/rpt_family="L2"	22849. .23100					
repeat_region	/rpt_family="MIR"	24663. .24827					
repeat_region	/rpt_family="AluSg"	25587. .25622					

Db 137936	AAACCTTGCAATTCCTGGGATAAAATATTACTTTGCTATGATGTGTAACATTTTGTATGT	137877
Qy	171 TCCT 174	
Db 137876	TGCT 137873	
RESULT 74		
AC123297		
LOCUS	Rattus norvegicus clone CH230-351P14, *** SEQUENCING IN PROGRESS	120714 bp DNA linear HTG 18-JUL-2002
DEFINITION	***, 53 unordered pieces.	
ACCESSION	AC123297	
VERSION	AC123297.2 GI:21746249	
KEYWORDS	HTG; HTGS PHASE1.	
SOURCE	Norway rat.	
ORGANISM	Rattus norvegicus	
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;	
	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;	
	Rattus.	
REFERENCE	1 (bases 1 to 120714)	
AUTHORS	Muzny,D.M., Adams,C., Adio-Oduola,B., Ali-osman,F.R., Allen,C., Alsbrooks,S.L., Amaratunge,H.C., Are,J.R., Ayele,M., Banks,T., Barbaria,J., Benton,J., Bimage,K., Blankenburg,K., Bonnin,D., Bouck,J., Bowie,S., Brieva,M., Brown,E., Brown,M., Bryant,N.P., Buhay,C., Burch,P., Burkett,C., Burrell,K.L., Byrd,N.C., Carron,T.F., Carter,M., Cavazos,S.R., Chacko,J., Chavez,D., Chen,G., Chen,R., Chen,Z., Chowdhry,I., Christopoulos,C., Cleveland,C.D., Cox,C., Coyle,M.D., Dathorne,S.R., David,R., Davila,M.L., Davis,C., Davy-Carroll,L., Dederich,D.A., Delaney,K.R., Delgado,O., Denn,A.L., Ding,Y., Dinh,H.H., Douthwaite,K.J., Draper,H., Dugan-Rocha,S., Durbin,K.J., Earnhart,C., Edgar,D., Edwards,C.C., Elhaj,C., Escotto,M., Falls,T., Ferraguto,D., Flagg,N., Ford,J., Foster,P., Frantz,P., Gabisi,A., Gao,J., Garcia,A., Garner,T., Garza,N., Gill,R., Gorrell,J.H., Guevara,W., Gunaratne,P., Hale,S., Hamilton,K., Harris,C., Harris,K., Hart,M., Havlak,P., Hawes,A., Hernandez,J., Hernandez,O., Hodgson,A., Hogue,M., Holloway,C., Hollins,B., Honsi,F., Howard,S., Huber,J., Hulyk,S., Hume,J., Jackson,L.E., Jacobson,B., Jia,Y., Johnson,R., Jolivet,S., Joudah,S., Karlsson,E., Kelly,S., Khan,U., King,L., Korvah,J., Kovar,C., Kratovic,J., Kureshi,A., Landry,N., Leal,B., Lewis,L.C., Lewis,L., Li,J., Li,Z., Lichtarge,O., Lieu,C., Liu,J., Liu,W., Loulseged,H., Lozado,R.J., Lu,X., Lucier,A., Lucier,R., Luna,R., Ma,J., Maheshwari,M., Mapua,P., Martin,R., Martindale,A., Martinez,E., Massey,E., Mawhiney,E., McLeod,M.P., Meador,M., Mei,G., Metzker,M., Miner,G., Miner,Z., Mitchell,T., Mohabbat,K., Morgan,M., Morris,S., Moser,M., Neal,D., Newton,J., Newton,N., Nguyen,A., Nguyen,N., Nguyen,N., Nickerson,E., Nwokenkwo,S., Oguh,M., Okwuonu,G., Oragunye,N., Oviedo,R., Pace,A., Payton,B., Peery,J., Perez,L., Peters,L., Pickens,R., Primus,E., Pu,L.L., Quiles,M., Ren,Y., Rives,M., Rojas,A., Rojubokan,I., Rolfe,M., Ruiz,S., Savery,G., Scherer,S., Scott,G., Shen,H., Shoostari,N., Sisson,I., Sodergren,E., Sonaike,T., Sparks,A., Stanley,H., Stone,H., Sutton,A., Svatek,A., Tabor,P., Tamerisa,A., Tamerisa,K., Tang,H., Tansey,J., Taylor,C., Taylor,T., Telford,B., Thomas,N., Thomas,S., Usmani,K., Vasquez,L., Vera,V., Villalon,D., Vinson,R., Wang,Q., Wang,S., Ward-Moore,S., Warren,R., Washington,C., Watlington,S., Williams,G., Williamson,A., Wleczyk,R., Wooden,S., Worley,K., Wu,C., Wu,Y., Wu,Y.F., Zhou,J., Zorrilla,S., Nelson,D., Weinstein,G. and Gibbs,R.	
	Direct Submission	
TITLE	Unpublished	
JOURNAL	2 (bases 1 to 120714)	
REFERENCE	Worley,K.C.	
AUTHORS	Direct Submission	
TITLE	Submitted (29-MAY-2002) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA	
JOURNAL	3 (bases 1 to 120714)	
REFERENCE	Worley,K.C.	
AUTHORS	Direct Submission	
TITLE	Submitted (18-JUL-2002) Human Genome Sequencing Center, Department	
JOURNAL		

Db 50980 ATGTACAGGTAATCTTACTGAAACCTAAGTCTTGATATGGTTCTAGTGTAAAGTCTGT 51039
QY 191 CGTTCCA 197
Db 51040 CGTTACA 51046

RESULT 75
AC117109
LOCUS
DEFINITION Rattus norvegicus clone CH230-242N20, *** SEQUENCING IN PROGRESS
***, 34 unordered pieces.
AC117109
AC117109.2 GI:21745874
VERSION HTG; HTGS PHASE1.
KEYWORDS Norway rat.
SOURCE Rattus norvegicus
ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
Rattus.
1 (bases 1 to 142597)
Muzny,D.M., Adams,C., Adio-Oduola,B., Ali-osman,F.R., Allen,C.,
Alsbrooks,S.L., Amaratunge,H.C., Are,J.R., Ayele,M., Banks,T.,
Barbaria,J., Benton,J., Bimage,K., Blankenburg,K., Bonnin,D.,
Bouck,J., Bowie,S., Brieva,M., Brown,E., Brown,M., Bryant,N.P.,
Buhay,C., Burch,P., Burkett,C., Burrell,K.L., Byrd,N.C.,
Carron,T.F., Carter,M., Cavazos,S.R., Chacko,J., Chavez,D.,
Chen,G., Chen,R., Chen,Z., Chowdhry,I., Christopoulos,C.,
Cleveland,C.D., Cox,C., Coyle,M.D., Dathorne,S.R., David,R.,
Davila,M.L., Davis,C., Davy-Carroll,L., Dederich,D.A.,
Delaney,K.R., Delgado,O., Denn,A.L., Ding,Y., Dinh,H.H.,
Douthwaite,K.J., Draper,H., Dugan-Rocha,S., Durbin,K.J.,
Earnhart,C., Edgar,D., Edwards,C.C., Elhaj,C., Escotto,M.,
Falls,T., Ferraguto,D., Flagg,N., Ford,J., Foster,P., Frantz,P.,
Gabisi,A., Gao,J., Garcia,A., Garner,T., Garza,N., Gill,R.,
Gorrell,J.H., Guevara,W., Gunaratne,P., Hale,S., Hamilton,K.,
Harris,C., Harris,K., Hart,M., Havlak,P., Hawes,A., Hernandez,J.,
Hernandez,O., Hodgson,A., Hogue,M., Holloway,C., Hollins,B.,
Homsy,F., Howard,S., Huber,J., Hulyk,S., Hume,J., Jackson,L.E.,
Jacobson,B., Jia,Y., Johnson,R., Jolivet,S., Joudah,S.,
Karlsson,E., Kelly,S., Khan,U., King,L., Korvah,J., Kovar,C.,
Kratovic,J., Kureshi,A., Landry,N., Leal,B., Lewis,L.C., Lewis,L.,
Li,J., Li,Z., Lichtarge,O., Liew,C., Liu,J., Liu,W., Loulseged,H.,
Lozado,R.J., Lu,X., Lucier,A., Lucier,R., Luna,R., Ma,J.,
Maheshwari,M., Mapua,P., Martin,R., Martindale,A., Martinez,E.,
Massey,E., Mawhiney,E., McLeod,M.P., Meador,M., Mei,G., Metzker,M.,
Miner,G., Miner,Z., Mitchell,T., Mohabbat,K., Morgan,M., Morris,S.,
Moser,M., Neal,D., Newton,J., Newton,N., Nguyen,A., Nguyen,N.,
Nguyen,N., Nickerson,E., Nwokenwo,S., Oguh,M., Okwuonu,G.,
Oragunye,N., Oviedo,R., Pace,A., Payton,B., Peery,J., Perez,L.,
Peters,L., Pickens,R., Primus,E., Pu,L.L., Quiles,M., Ren,Y.,
Rives,M., Rojas,A., Rojibokan,I., Rolfe,M., Ruiz,S., Savary,G.,
Scherer,S., Scott,G., Shen,H., Shooshtari,N., Sisson,I.,
Sodergren,E., Sonaike,T., Sparks,A., Stanley,H., Stone,H.,
Sutton,A., Svatek,A., Tabor,P., Tamerisa,A., Tamerisa,K., Tang,H.,
Tansey,J., Taylor,C., Taylor,T., Telford,B., Thomas,N., Thomas,S.,
Usmani,K., Vasquez,L., Vera,V., Villalon,D., Vinson,R., Wang,Q.,
Wang,S., Ward-Moore,S., Warren,R., Washington,C., Watlington,S.,
Williams,G., Williamson,A., Wleczyk,R., Wooden,S., Worley,K.,
Wu,C., Wu,Y., Wu,Y.F., Zhou,J., Zorrilla,S., Nelson,D.,
Weinstock,G. and Gibbs,R.
Direct Submission
Unpublished
2 (bases 1 to 142597)
Worley,K.C.
Direct Submission
TITLE
Submitted (08-APR-2002) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA
3 (bases 1 to 142597)
Worley,K.C.
Direct Submission
TITLE

REFERENCE
AUTHORS
TITLE
Submitted (18-JUL-2002) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA
On Jul 14, 2002 this sequence version replaced gi:20069615.
----- Genome Center
Center: Baylor College of Medicine
Center code: BCM
Web site: <http://www.hgsc.bcm.tmc.edu/>
Contact: hgsc-help@bcm.tmc.edu
----- Project Information
Center project name: GTNB
Center clone name: CH230-242N20
----- Summary Statistics
Sequencing vector: Plasmid;
Chemistry: Dye-terminator Big Dye; 100% of reads
Assembly program: Phrap; version 0.990329
Consensus quality: 117259 bases at least Q40
Consensus quality: 120122 bases at least Q30
Consensus quality: 122747 bases at least Q20

* NOTE: Estimated insert size may differ from sequence length
* (see http://www.hgsc.bcm.tmc.edu/docs/Genbank_draft_data.html).
* NOTE: This is a 'working draft' sequence. It currently
* consists of 34 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
* 1 1769: contig of 1769 bp in length
* 1770 1869: gap of unknown length
* 1870 2963: contig of 1094 bp in length
* 2964 3063: gap of unknown length
* 3064 4456: contig of 1393 bp in length
* 4457 4556: gap of unknown length
* 4557 5992: contig of 1436 bp in length
* 5993 6092: gap of unknown length
* 6093 7489: contig of 1397 bp in length
* 7490 7589: gap of unknown length
* 7590 9173: contig of 1584 bp in length
* 9174 9273: gap of unknown length
* 9274 11148: contig of 1875 bp in length
* 11149 11248: gap of unknown length
* 11249 13917: contig of 2669 bp in length
* 13918 14017: gap of unknown length
* 14018 16203: contig of 2186 bp in length
* 16204 16303: gap of unknown length
* 16304 17917: contig of 1614 bp in length
* 17918 18017: gap of unknown length
* 18018 19705: contig of 1688 bp in length
* 19706 19805: gap of unknown length
* 19806 22857: contig of 3052 bp in length
* 22858 22957: gap of unknown length
* 22958 25118: contig of 2161 bp in length
* 25119 25218: gap of unknown length
* 25219 27161: contig of 1943 bp in length
* 27162 27261: gap of unknown length
* 27262 29591: contig of 2330 bp in length
* 29592 29691: gap of unknown length
* 29692 30909: contig of 1218 bp in length
* 30910 31009: gap of unknown length
* 31010 33313: contig of 2304 bp in length
* 33314 33413: gap of unknown length
* 33414 36414: contig of 3001 bp in length
* 36415 36514: gap of unknown length
* 36515 41632: contig of 5118 bp in length
* 41633 41732: gap of unknown length
* 41733 45404: contig of 3672 bp in length
* 45405 45504: gap of unknown length
* 45505 50390: contig of 4886 bp in length
* 50391 50490: gap of unknown length
* 50491 54730: contig of 4240 bp in length
* *

Search completed: February 15, 2003, 22:44:04
Job time : 1527.5 secs

GenCore version 5.1.4 p5 4578
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: February 15, 2003, 20:20:11 ; Search time 66 Seconds
(without alignments)
264.857 Million cell updates/sec

Title: 09-833799-13B
Perfect score: 327
Sequence: 1 aqepvkgpvstkgpscpil.....cpgikkcecgscgmcfvpq 57

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 441362 seqs, 153338381 residues

Total number of hits satisfying chosen parameters: 882724

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 100 summaries

Command line parameters:
-MODEL=frame+p2n.model -DEV=xlp
-O=/cgn2_1/USPTO.spool/US09833799/runat 12022003 120043 23688/app.query.fasta_1.199
-DB=Issued Patents_NA -QFMT=fastap -SUFFIX=rni -MINMATCH=0.1 -LOOPCL=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi
-LIST=100 -DOCALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=75
-MODE=LOCAL -OUTFMT=ptc -HEAPSIZ=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US09833799@cgn1 1 61 @runat 12022003 120043 23688 -NCPU=6 -ICPU=3
-NO_XLPXY -NO_MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -LONGLOG -DEV_TIMEOUT=120
-WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Issued Patents_NA:
1: /cgn2_6/ptodata/1/ina/5A_COMB.seq:*
2: /cgn2_6/ptodata/1/ina/5B_COMB.seq:*
3: /cgn2_6/ptodata/1/ina/6A_COMB.seq:*
4: /cgn2_6/ptodata/1/ina/6B_COMB.seq:*
5: /cgn2_6/ptodata/1/ina/PTUS_COMB.seq:*
6: /cgn2_6/ptodata/1/ina/backfiles1.seq:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	327	100.0	571	1	US-08-322-742-14
2	324	99.1	177	1	US-08-379-437-3
3	323	98.8	177	1	US-08-379-437-5
4	323	98.8	177	1	US-08-379-437-7
5	322	98.5	321	1	US-08-322-742-11
6	143	43.7	194	2	US-07-963-538B-5
7	139	42.5	180	3	US-08-483-503A-2
8	105	32.1	292	4	US-09-404-879A-377
9	98	30.0	724	4	US-09-724-864-23
10	94.5	28.9	437	4	US-09-383-586-29
11	92.5	28.3	6314	1	US-08-211-430-1
12	85	26.0	753	4	US-09-152-060-34
					Sequence 14, Appl
					Sequence 3, Appli
					Sequence 5, Appli
					Sequence 7, Appli
					Sequence 11, Appl
					Sequence 5, Appli
					Sequence 2, Appli
					Sequence 377, App
					Sequence 23, Appl
					Sequence 29, Appl
					Sequence 1, Appli
					Sequence 34, Appl

13	85	26.0	783	4	US-09-152-060-50	Sequence 50, Appl
14	81.5	24.9	1042	2	US-08-761-248B-5	Sequence 5, Appli
15	81.5	24.9	1071	2	US-08-761-248B-1	Sequence 1, Appli
16	79.5	24.3	1366	2	US-08-761-248B-3	Sequence 3, Appli
17	77.5	23.7	4157	2	US-08-162-146-2	Sequence 2, Appli
18	77.5	23.7	4157	4	US-09-314-127-2	Sequence 2, Appli
19	75	22.9	80	1	US-08-304-051-7	Sequence 7, Appli
20	75	22.9	80	1	US-08-304-051-11	Sequence 11, Appl
21	75	22.9	80	5	PCT-US95-11445-7	Sequence 7, Appli
22	75	22.9	80	5	PCT-US95-11445-11	Sequence 11, Appl
23	69	21.1	403	4	US-09-404-879A-198	Sequence 198, App
24	69	21.1	2065	3	US-08-335-865J-8	Sequence 8, Appli
25	68	20.8	1419	4	US-09-434-288-2	Sequence 2, Appli
26	66.5	20.3	11444	2	US-08-222-617A-26	Sequence 26, Appl
27	66.5	20.3	12364	2	US-08-222-617A-1	Sequence 1, Appli
28	65	19.9	68	2	US-07-963-538B-34	Sequence 34, Appl
29	64.5	19.7	1364	1	US-08-306-691B-50	Sequence 50, Appl
30	64.5	19.7	1364	5	PCT-US93-06251-65	Sequence 65, Appl
31	64.5	19.7	1590	4	US-08-961-527-184	Sequence 184, App
32	64.5	19.7	2625	2	US-08-357-533A-1	Sequence 1, Appli
33	64.5	19.7	2625	2	US-08-459-009-1	Sequence 1, Appli
34	64.5	19.7	2625	3	US-08-459-951-1	Sequence 1, Appli
35	64	19.6	80161	3	US-09-036-987A-1	Sequence 1, Appli
36	64	19.6	80161	4	US-09-370-700-1	Sequence 1, Appli
37	63.5	19.4	4170	2	US-08-619-198-4	Sequence 4, Appli
38	63.5	19.4	5192	2	US-08-619-198-8	Sequence 8, Appli
39	63.5	19.4	5198	2	US-08-619-198-1	Sequence 1, Appli
40	63	19.3	8931	3	US-09-028-934-28	Sequence 28, Appl
41	62.5	19.1	3855	4	US-08-974-549A-4	Sequence 4, Appli
42	62.5	19.1	4015	3	US-08-851-843A-224	Sequence 224, App
43	62.5	19.1	4015	4	US-08-974-549A-1	Sequence 1, Appli
44	62.5	19.1	4015	4	US-08-854-050-224	Sequence 224, App
45	62.5	19.1	4015	4	US-09-430-323-224	Sequence 224, App
46	62.5	19.1	4015	4	US-09-572-423B-3	Sequence 3, Appli
47	62.5	19.1	4015	4	US-09-128-354-1	Sequence 1, Appli
48	62.5	19.1	4015	4	US-09-675-321-1	Sequence 1, Appli
49	62.5	19.1	4015	4	US-09-052-919-1	Sequence 1, Appli
50	62.5	19.1	4037	4	US-08-974-549A-343	Sequence 343, App
51	62.5	19.1	4335	4	US-08-974-549A-6	Sequence 6, Appli
52	62.5	19.1	6677	4	US-08-939-366-27	Sequence 27, Appl
53	62.5	19.1	6677	4	US-09-467-997-6	Sequence 6, Appli
54	61.5	18.8	708	2	US-08-400-159-3	Sequence 3, Appli
55	61.5	18.8	708	3	US-08-611-729A-3	Sequence 3, Appli
56	61.5	18.8	1665	2	US-08-886-152-2	Sequence 2, Appli
57	61.5	18.8	1665	4	US-09-196-222-2	Sequence 2, Appli
58	61.5	18.8	2892	1	US-08-264-534-5	Sequence 5, Appli
59	61.5	18.8	2892	1	US-08-083-590A-1	Sequence 1, Appli
60	61.5	18.8	2892	1	US-08-465-500-5	Sequence 5, Appli
61	61.5	18.8	2892	2	US-08-346-126-5	Sequence 5, Appli
62	61.5	18.8	2892	2	US-08-346-128-5	Sequence 5, Appli
63	61.5	18.8	2892	3	US-08-532-384-1	Sequence 1, Appli
64	61.5	18.8	2892	3	US-08-893-828-5	Sequence 5, Appli
65	61	18.7	631	2	US-08-747-887-4	Sequence 4, Appli
66	61	18.7	631	2	US-08-728-323A-4	Sequence 4, Appli
67	61	18.7	653	4	US-08-998-416-149	Sequence 149, App
68	61	18.7	1055	4	US-09-056-556-191	Sequence 191, App
69	61	18.7	1055	4	US-09-072-596-186	Sequence 186, App
70	60	18.3	518	4	US-09-615-192A-172	Sequence 172, App
71	60	18.3	850	4	US-09-449-285A-8	Sequence 8, Appli
72	60	18.3	1089	4	US-08-858-207A-161	Sequence 161, App
73	60	18.3	2783	4	US-09-282-147-37	Sequence 37, Appl
74	60	18.3	3222	4	US-09-543-084A-29	Sequence 29, Appl
75	60	18.3	4050	4	US-09-543-084A-26	Sequence 26, Appl
76	60	18.3	4093	4	US-09-543-084A-28	Sequence 28, Appl
77	60	18.3	8802	3	US-08-896-449A-1	Sequence 1, Appli
78	60	18.3	8802	3	US-09-132-652-1	Sequence 1, Appli
79	60	18.3	4403765	4	US-09-103-840A-2	Sequence 2, Appli
80	60	18.3	4411529	4	US-09-103-840A-1	Sequence 1, Appli
81	59.5	18.2	703	4	US-09-452-817-5	Sequence 5, Appli
82	59.5	18.2	717	4	US-08-998-416-849	Sequence 849, App
83	59.5	18.2	780	4	US-09-452-817-7	Sequence 7, Appli
84	59	18.0	67	2	US-07-963-538B-33	Sequence 33, Appl
85	59	18.0	610	4	US-09-072-596-293	Sequence 293, App

Query Match: 99.08% Indels: 0
DB: 1 Gaps: 0
09-833799-13B (1-57) x US-08-379-437-3 (1-177)
QY 1 AlaGlnGluProValLysGlyProValSerThrLysProGlySerCysProIleLeu 20
Db 1 GCACAGGAACCAAGTAAAGGTCCGGTGTGACCAAAACCGGGCTCTTGCCTGATTATCCTG 60
QY 21 IleArgCysAlaMetLeuAsnProProAsnArgCysLeuLysAspThrAspCysProGly 40
Db 61 ATCCGCTGCGCTTGTCTGAACCCCGGACCGGTGTCTGAAAGACACTGACTGCCCGGGT 120
QY 41 IleLysLysCysGluGlySerCysGlyMetAlaCysPheValProGln 57
Db 121 ATCAAAAATGCTGGAAGGTTCTTTCGGTATGGCATGCTTCGTTCCGCAG 171
RESULT 3
US-08-379-437-5
; Sequence 5, Application US/08379437
; Patent No. 5734014
; GENERAL INFORMATION:
; APPLICANT: ISHIMA, YOSHIKI
; APPLICANT: OKAWA, NORIYUKI
; APPLICANT: YOSHIDA, MASAYA
; APPLICANT: AMAGAYA, SAKAE
; APPLICANT: KAJI, AKIRA
; TITLE OF INVENTION: NOVEL ELAFIN DERIVATIVE
; NUMBER OF SEQUENCES: 16
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: OBLON, SPIVAK, MCCLELLAND, MAIER & NEUSTADT
; STREET: 1755 S. JEFFERSON DAVIS HWY, SUITE 400
; CITY: ARLINGTON
; STATE: VA
; COUNTRY: USA
; ZIP: 22202
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/379,437
; FILING DATE: 27-MAR-1995
; CLASSIFICATION: 530
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JP PCT/JP93/01133
; FILING DATE: 11-AUG-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JP 234085/1992
; FILING DATE: 11-AUG-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: OBLON, NORMAN F
; REGISTRATION NUMBER: 24,618
; TELEPHONE: 703-413-3000
; TELEFAX: 703-413-2220
; TELEX: 248855 OPAT UR
; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 177 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: other nucleic acid
; DESCRIPTION: /desc = "synthetic DNA"
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 1..171
US-08-379-437-5

Alignment Scores: 1.41e-33 Length: 177
Pred. No.: 1

Score: 323.00 Matches: 56
Percent Similarity: 100.00% Conservative: 1
Best Local Similarity: 98.25% Mismatches: 0
Query Match: 98.78% Indels: 0
DB: 1 Gaps: 0
09-833799-13B (1-57) x US-08-379-437-5 (1-177)
QY 1 AlaGlnGluProValLysGlyProValSerThrLysProGlySerCysProIleLeu 20
Db 1 GCACAGGAACCAAGTAAAGGTCCGGTGTGACCAAAACCGGGCTCTTGCCTGATTATCCTG 60
QY 21 IleArgCysAlaMetLeuAsnProProAsnArgCysLeuLysAspThrAspCysProGly 40
Db 61 ATCCGCTGCGCTTGTCTGAACCCCGGACCGGTGTCTGAAAGACACTGACTGCCCGGGT 120
QY 41 IleLysLysCysGluGlySerCysGlyMetAlaCysPheValProGln 57
Db 121 ATCAAAAATGCTGGAAGGTTCTTTCGGTATGGCATGCTTCGTTCCGCAG 171
RESULT 4
US-08-379-437-7
; Sequence 7, Application US/08379437
; Patent No. 5734014
; GENERAL INFORMATION:
; APPLICANT: ISHIMA, YOSHIKI
; APPLICANT: OKAWA, NORIYUKI
; APPLICANT: YOSHIDA, MASAYA
; APPLICANT: AMAGAYA, SAKAE
; APPLICANT: KAJI, AKIRA
; TITLE OF INVENTION: NOVEL ELAFIN DERIVATIVE
; NUMBER OF SEQUENCES: 16
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: OBLON, SPIVAK, MCCLELLAND, MAIER & NEUSTADT
; STREET: 1755 S. JEFFERSON DAVIS HWY, SUITE 400
; CITY: ARLINGTON
; STATE: VA
; COUNTRY: USA
; ZIP: 22202
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/379,437
; FILING DATE: 27-MAR-1995
; CLASSIFICATION: 530
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JP PCT/JP93/01133
; FILING DATE: 11-AUG-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JP 234085/1992
; FILING DATE: 11-AUG-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: OBLON, NORMAN F
; REGISTRATION NUMBER: 24,618
; TELEPHONE: 703-413-3000
; TELEFAX: 703-413-2220
; TELEX: 248855 OPAT UR
; INFORMATION FOR SEQ ID NO: 7:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 177 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: other nucleic acid
; DESCRIPTION: /desc = "synthetic DNA"
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 1..171
US-08-379-437-7

; REGISTRATION NUMBER: 30,092
; REFERENCE/DOCKET NUMBER: TEJN-005/02US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-843-5070
; TELEFAX: 415-857-0663
; TELEX: 380816 COOLEY PA
; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 194 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (synthetic)
US-07-963-538B-5

Alignment Scores:
Pred. No.: 2.14e-10 Length: 194
Score: 143.00 Matches: 27
Percent Similarity: 55.36% Conservative: 4
Best Local Similarity: 48.21% Mismatches: 23
Query Match: 43.73% Indels: 2
DB: 2 Gaps: 1

09-833799-13B (1-57) x US-07-963-538B-5 (1-194)

QY 3 GluProValLysGlyProValSerThr-----LysProGlySerCysProIleLeu 20
Db 1 GATCCGGTCGACACCCCGAACCCGACGCGTCGTAAACCGGTAATGTCGGTTACATAT 60
QY 21 IleArgCysAlaMetLeuAsnProProAsnArgCysLeuLysAspThrAspCysProGly 40
Db 61 GGTCAGTGTCTGATGCTGAACCCGCCGAACTTCTGTGAAATGGACGTCAGTGTAACGA 120
QY 41 IleLysLysCysCysGluGlySerCysGlyMetAlaCysPheValPro 56
Db 121 GATCTGAAATGTTGATGGGTATGTGTGGTAAATCTTGTGTTTCTCCG 168

RESULT 7
US-08-483-503A-2
; Sequence 2, Application US/08483503A
; Patent No. 6017880
; GENERAL INFORMATION:
; APPLICANT: Eisenberg, Stephen
; APPLICANT: Wahl, Sharon M.
; APPLICANT: Thompson, Robert C.
; APPLICANT: Dripps, David J.
; TITLE OF INVENTION: INHIBITION OF RETROVIRUS INFECTION
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Finnegan, Henderson, Farabow, Garrett &
; ADDRESSEE: Dunner, L.L.P.
; STREET: 1300 I Street, N.W.
; CITY: Washington
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20005-3315
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/483,503A
; FILING DATE: 07-JUN-1995
; CLASSIFICATION: 514
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/209,040
; FILING DATE: 09-MAR-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/943,369
; FILING DATE: 09-SEP-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: Barker, M. Paul

; REGISTRATION NUMBER: 32,013
; REFERENCE/DOCKET NUMBER: 04189.0084-02000
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 202-408-4000
; TELEFAX: 202-408-4400
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 180 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
US-08-483-503A-2

Alignment Scores:
Pred. No.: 6.32e-10 Length: 180
Score: 139.00 Matches: 26
Percent Similarity: 53.57% Conservative: 4
Best Local Similarity: 46.43% Mismatches: 24
Query Match: 42.51% Indels: 2
DB: 3 Gaps: 1

09-833799-13B (1-57) x US-08-483-503A-2 (1-180)

QY 3 GluProValLysGlyProValSerThr-----LysProGlySerCysProIleLeu 20
Db 4 GATCCTGTTGACACCCCAACACCAAGAGGAGCCTGGGAAGTGGCCAGTGACTTAT 63
QY 21 IleArgCysAlaMetLeuAsnProProAsnArgCysLeuLysAspThrAspCysProGly 40
Db 64 GGCCAATGTTGATGCTTAACCCCCCAATTTCTGTGAGATGGATGGCCAGTCAAGCGT 123
QY 41 IleLysLysCysCysGluGlySerCysGlyMetAlaCysPheValPro 56
Db 124 GACTTGAAGTGTTCATGGGCATGTGTGGGAAATCCTGCGTTTCCCCT 171

RESULT 8
US-09-404-879A-377
; Sequence 377, Application US/09404879A
; Patent No. 6468546
; GENERAL INFORMATION:
; APPLICANT: Mitcham, Jennifer L.
; APPLICANT: King, Gordon E.
; APPLICANT: Algate, Paul A.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
; TITLE OF INVENTION: DIAGNOSIS OF OVARIAN CANCER
; FILE REFERENCE: 210121.462C2
; CURRENT APPLICATION NUMBER: US/09/404,879A
; NUMBER OF SEQ ID NOS: 393
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 377
; LENGTH: 292
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-404-879A-377

Alignment Scores:
Pred. No.: 2.84e-05 Length: 292
Score: 105.00 Matches: 18
Percent Similarity: 57.14% Conservative: 2
Best Local Similarity: 51.43% Mismatches: 15
Query Match: 32.11% Indels: 0
DB: 4 Gaps: 0

09-833799-13B (1-57) x US-09-404-879A-377 (1-292)

QY 22 ArgCysAlaMetLeuAsnProProAsnArgCysLeuLysAspThrAspCysProGlyIle 41
Db 1 CAATGTTTGATGCTTAACCCCCCAATTTCTGTGAGATGGATGGCCAGTCAAGCGTGAC 60
QY 42 LysLysCysCysGluGlySerCysGlyMetAlaCysPheValPro 56

Db 61 TTGAAGTGTTCATGGGCATGTGTGGGAAATCCTGCGTTTCCCCT 105

RESULT 9

US-09-724-864-23

; Sequence 23, Application US/09724864

; Patent No. 6380362

; GENERAL INFORMATION:

; APPLICANT: Watson, James D

; APPLICANT: Murison, James G.

; TITLE OF INVENTION: Polynucleotides, polypeptides expressed

; TITLE OF INVENTION: by the polynucleotides and methods for their use.

; FILE REFERENCE: 11000.1050U1

; CURRENT APPLICATION NUMBER: US/09/724,864

; CURRENT FILING DATE: 2000-11-28

; PRIOR APPLICATION NUMBER: U.S. No. 6380362 60/171,678

; PRIOR FILING DATE: 1999-12-23

; NUMBER OF SEQ ID NOS: 72

; SOFTWARE: FastSEQ for Windows Version 4.0

; SEQ ID NO 23

; LENGTH: 724

; TYPE: DNA

; ORGANISM: Mouse

US-09-724-864-23

Alignment Scores:

Pred. No.:	0.000767	Length:	724
Score:	98.00	Matches:	25
Percent Similarity:	43.24%	Conservative:	7
Best Local Similarity:	33.78%	Mismatches:	18
Query Match:	29.97%	Indels:	24
DB:	4	Gaps:	4

09-833799-13B (1-57) x US-09-724-864-23 (1-724)

QY 3 GluProValLysGlyProValSerThrLysPro-----13

Db 313 AAACCACCGGAGGTCAAGTCTCCACGAGCCACCGGTGTGACGAGGAAGGCTTAGGT 372

QY 14 -----GlySerCysProIleIle-----LeuIleArgCysAla 24

Db 373 GTCCGAGAAAGCAGGCGACCTGCCCGGCGGTGGACATACCCAGCTCGGCCTCTGT--- 429

QY 25 MetLeuAsnProProAsnArgCysLeuLysAspThrAspCysProGlyIleLysLysCys 44

Db 430 -----GAGGACCAAGTGTGAGGTGACAGCCAGTGTCTTGGCAACATGAATGC 477

QY 45 CysGluGlySerCysGly---MetAlaCysPheValProGln 57

Db 478 TGCCGCAATGGATGTGGGAAGATGGCCTGCACCACACCCAAA 519

RESULT 10

US-09-383-586-29

; Sequence 29, Application US/09383586

; Patent No. 6242419

; GENERAL INFORMATION:

; APPLICANT: Strachan, Lorna

; APPLICANT: Sleeman, Matthew

; APPLICANT: Abernethy, Nevin

; APPLICANT: Onrust, Rene

; APPLICANT: Kumble, Anand

; APPLICANT: Murison, Greg

; TITLE OF INVENTION: Compounds isolated from stromal cells

; TITLE OF INVENTION: and methods for their use

; FILE REFERENCE: 11000.1037c1

; CURRENT APPLICATION NUMBER: US/09/383,586

; CURRENT FILING DATE: 1999-08-26

; NUMBER OF SEQ ID NOS: 38

; SOFTWARE: FastSEQ for Windows Version 3.0

; SEQ ID NO 29

; LENGTH: 437

; TYPE: DNA

; ORGANISM: Mouse

US-09-383-586-29

Alignment Scores:

Pred. No.:	0.00109	Length:	437
Score:	94.50	Matches:	19
Percent Similarity:	41.51%	Conservative:	3
Best Local Similarity:	35.85%	Mismatches:	14
Query Match:	28.90%	Indels:	17
DB:	4	Gaps:	2

09-833799-13B (1-57) x US-09-383-586-29 (1-437)

QY 12 LysProGlySerCysProIleIleLeuIleArgCysAlaMetLeuAsnProPro-----29

Db 80 AAACCTGGAGCTGTCCC-----AAGCCTTCACCAGAA 112

QY 30 -----AsnArgCysLeuLysAspThrAspCysProGlyIleLysLys 43

Db 113 AGTGTGGAAATTGTGTTGATCAATGCTCAGGAGATGGATCCTGCGCCTGGCAACATGAAG 172

QY 44 CysCysGluGlySerCysGlyMetAlaCysPheValPro 56

Db 173 TGCTGTAGCAATAGCTGTGGTCATGTCGCAAAACTCCT 211

RESULT 11

US-08-211-430-1

; Sequence 1, Application US/08211430

; Patent No. 5763166

; GENERAL INFORMATION:

; APPLICANT:

; TITLE OF INVENTION: NUCLEIC SEQUENCE OF THE GENE ASSOCIATED WITH

; TITLE OF INVENTION: X CHROMOSOME LINKED KALLMANN SYNDROME, CORRESPONDING

; TITLE OF INVENTION: PEPTIDE SEQUENCES, DIAGNOSTIC APPLICATIONS.

; NUMBER OF SEQUENCES: 32

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: PatentIn Release #1.0, Version #1.25 (EPO)

; CURRENT APPLICATION DATA: US/08/211,430

; APPLICATION FOR SEQ ID NO: 1:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 6314 base pairs

; TYPE: nucleic acid

; STRANDEDNESS: double

; TOPOLOGY: linear

; MOLECULE TYPE: cDNA

; HYPOTHETICAL: NO

; ANTI-SENSE: NO

; ORIGINAL SOURCE:

; ORGANISM: Homo sapiens

; CELL LINE: foetal brain cells

US-08-211-430-1

Alignment Scores:

Pred. No.:	0.0726	Length:	6314
Score:	92.50	Matches:	18
Percent Similarity:	40.38%	Conservative:	3
Best Local Similarity:	34.62%	Mismatches:	20
Query Match:	28.29%	Indels:	11
DB:	1	Gaps:	2

09-833799-13B (1-57) x US-08-211-430-1 (1-6314)

QY 12 LysProGlySerCysProIle-----IleLeuIleArgCysAlaMet 25

Db 538 AAGCAGGGGACTGTCCGGCTCCTGAGAAAGCCAGTGGATTTCGGCGCCCTGTGTT--- 594

QY 26 LeuAsnProProAsnArgCysLeuLysAspThrAspCysProGlyIleLysLysCys 45

Db 595 -----GAAAGCTGCGAAGTTGACATGAGTGTCTCTGGGTGAAGAAATGTTGT 642

QY 46 GluGlySerCysGlyMetAlaCysPheValProGln 57

Db 2938 TCCCAGTTGTGTGAGGAGCTGAGCGACTGTGCCAACGACATCGAGTGCAGGGGCGACAAG 2997

Qy 43 LysCysCysGluGlySerCysGlyMetAlaCysPheValPro 56

; INDEX: 304138
; INFORMATION FOR SEQ ID NO: 2:

TEL: 904136

; INDEX: 204138
; INFORMATION FOR SEQ ID NO: 2:

```

; INFORMATION FOR SEQ ID NO: 7:
;
; SEQUENCE CHARACTERISTICS:
;
; LENGTH: 80 bases
;
; TYPE: nucleic acid
;
; STRANDEDNESS: single

```

US-08-304-051-11

09-833799-13B (1-57) x PCT-US95-11445-7 (1-80)

Patent No. 6468546

```
; GENERAL INFORMATION:
; APPLICANT: Mitcham, Jennifer L.
; APPLICANT: King, Gordon E.
; APPLICANT: Algate, Paul A.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
; TITLE OF INVENTION: DIAGNOSIS OF OVARIAN CANCER
; FILE REFERENCE: 210121.462C2
; CURRENT APPLICATION NUMBER: US/09/404,879A
; CURRENT FILING DATE: 1999-09-24
; NUMBER OF SEQ ID NOS: 393
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 198
; LENGTH: 403
; TYPE: DNA
; ORGANISM: Homo sapien
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)..(403)
; OTHER INFORMATION: n = A,T,C or G
US-09-404-879A-198

Alignment Scores:
Pred. No.: 1.85 Length: 403
Score: 69.00 Matches: 18
Percent Similarity: 50.00% Conservatives: 6
Best Local Similarity: 37.50% Mismatches: 21
Query Match: 21.10% Indels: 3
DB: 4 Gaps: 2

09-833799-13B (1-57) x US-09-404-879A-198 (1-403)
QY 12 LysProGlySerCysProIleIleLeuIleArgCysAlaMetLeuAsn---ProPro-As 30
Db 342 AAGAGGGTTCCTGCCCCCAGGTGAACATAACTTCCCCAGCTCGGGCTTCTGTGGGA 283
QY 30 nArgCysLeuLysAspThrAspCysProGlyIleLysLysCysCysGluGlySerCysG1 50
Db 282 CCAGTGCCAGGTGACAGCCAGTGCTCTGCCCAGATGAATGCCGCCGAATGGCTGTGG 223
QY 50 y---MetAlaCysPheValPro 56
Db 222 GAAGGTGCTCTGTCTCACTCCC 201

RESULT 24
US-08-335-865J-8
; Sequence 8, Application US/08335865J
; Patent No. 6107472
; GENERAL INFORMATION:
; APPLICANT: Stackner, Steven A.; Hovens, Christopher M.,
; APPLICANT: Wilks, Andrew F.
; TITLE OF INVENTION: RECEPTOR-TYPE TYROSINE KINASE-LIKE MOLECULES
; NUMBER OF SEQUENCES: 21
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fulbright & Jaworski L.L.P.
; STREET: 666 Fifth Ave
; CITY: New York
; STATE: New York
; COUNTRY: USA
; ZIP: 10103
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.5 inch, 720 KB storage
; COMPUTER: IBM PS/2
; OPERATING SYSTEM: PC-DOS
; SOFTWARE: ASCII/Wordperfect
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/335,865J
; FILING DATE: 19-January-1995
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/AU93/00210
; FILING DATE: 10-May-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PL2358
```

```
; FILING DATE: 11-May-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: Hanson, No. 6107472man D.
; REGISTRATION NUMBER: 30,946
; REFERENCE/DOCKET NUMBER: LUD-5277
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 318-3100
; TELEFAX: (212) 752-5958
; INFORMATION FOR SEQ ID NO: 8:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2065
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
US-08-335-865J-8

Alignment Scores:
Pred. No.: 16.8 Length: 2065
Score: 69.00 Matches: 16
Percent Similarity: 47.83% Conservatives: 6
Best Local Similarity: 34.78% Mismatches: 22
Query Match: 21.10% Indels: 2
DB: 3 Gaps: 1

09-833799-13B (1-57) x US-08-335-865J-8 (1-2065)
QY 4 ProValLysGlyProValSerThrLysProGlySerCysProIleIleLeuIleArgCys 23
Db 1828 CCAGCAGCTGGTCCAGTGCCTCACAGAGTTCACAGCTGCCCTGGGAGCCTACGTCGACT 1887
QY 24 AlaMetLeuAsnProPro-----AsnArgCysLeuLysAspThrAspCysProGlyIle 41
Db 1888 TCTCTCCCCATGCCGCCACTCAGAAAGTGCCTGTCTGTACAGGATGCCCTCGTGCA 1947
QY 42 LysLysCysCysGluGly 47
Db 1948 GCGCAGTGCCTGCAGGGG 1965

RESULT 25
US-09-434-288-2
; Sequence 2, Application US/09434288
; Patent No. 6303767
; GENERAL INFORMATION:
; APPLICANT: Betlach C., Melanie
; APPLICANT: McDaniel, Robert
; TITLE OF INVENTION: POLYKETIDE SYNTHASE ENZYMES AND RECOMBINANT DNA
; TITLE OF INVENTION: CONSTRUCTS THEREFOR
; FILE REFERENCE: 30062-20030.00
; CURRENT APPLICATION NUMBER: US/09/434,288
; CURRENT FILING DATE: 1999-11-05
; PRIOR APPLICATION NUMBER: 60/107,093
; PRIOR FILING DATE: 1998-11-05
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
; LENGTH: 1419
; TYPE: DNA
; ORGANISM: Streptomyces narbonensis
US-09-434-288-2

Alignment Scores:
Pred. No.: 13.6 Length: 1419
Score: 68.00 Matches: 18
Percent Similarity: 44.44% Conservatives: 6
Best Local Similarity: 33.33% Mismatches: 14
Query Match: 20.80% Indels: 16
DB: 4 Gaps: 4

09-833799-13B (1-57) x US-09-434-288-2 (1-1419)
QY 13 ProGlySerCysProIleIleLeuIleArg---CysAlaMetLeuAsnProProAsnArg 31
Db 786 CCCCCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 845
```

QY 32 -----CysLeuLysAspThrAsp---CysProGlyIleLys----- 42

Db 846 GTCCCCGTGCGGTGCGCTCGACGACATCGACCTCGCCCCGGGTGAGCCGGCGTCCGCC 905

QY 43 -----LysCysCysGluGlySerCysGly 50

Db 906 AGGGCCCGCGGATGACGCGCTGCTGCGACGGCCGCTTCGGA 947

RESULT 26

US-08-222-617A-26/c

; Sequence 26, Application US/08222617A

; Patent No. 5882879

; GENERAL INFORMATION:

; APPLICANT: Veenstra, Annemarie E.

; APPLICANT: Martin, Juan F.

; APPLICANT: Garcia, Bruno D.

; APPLICANT: Gutierrez, Santiago

; APPLICANT: Barredo, Jose L.

; APPLICANT: Von Doehren, Hans

; APPLICANT: Palissa, Harriet

; APPLICANT: Van Liempt, Henk

; APPLICANT: Montenegro, Eduardo P.

; TITLE OF INVENTION: A Method for Influencing Beta-Lactam

; TITLE OF INVENTION: Antibiotic Production and for Isolation of Large

; TITLE OF INVENTION: Quantities of ACV Synthetase

; NUMBER OF SEQUENCES: 27

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: McDonnell Boehnen Hulbert & Berghoff

; STREET: 300 South Wacker Drive

; CITY: Chicago

; STATE: Illinois

; COUNTRY: USA

; ZIP: 60606

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: PatentIn Release #1.0, Version #1.25

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/222,617A

; FILING DATE: 04-APR-1994

; CLASSIFICATION: 435

; ATTORNEY/AGENT INFORMATION:

; REFERENCE/DOCKET NUMBER: 97,157

; INFORMATION FOR SEQ ID NO: 26:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 11444 base pairs

; TYPE: nucleic acid

; STRANDEDNESS: single

; TOPOLOGY: linear

; MOLECULE TYPE: DNA (genomic)

US-08-222-617A-26

Alignment Scores:

Pred. No.: 355 Length: 11444

Score: 66.50 Matches: 17

Percent Similarity: 35.44% Conservative: 11

Best Local Similarity: 21.52% Mismatches: 26

Query Match: 20.34% Indels: 25

DB: 2 Gaps: 2

09-833799-13B (1-57) x US-08-222-617A-26 (1-11444)

QY 2 GlnGluProValLysGlyProValSerThrLysProGlySerCysProIleIle----- 19

Db 5877 CAACCAGGAATTCTGGGCCCCCTTCTCTGCAATCCTTGGCAATCACCACAGACTGTTGA 5818

QY 20 -----LeuIleArgCysAlaMetLeuAsn 27

Db 5817 TTCCGTGATAGACGATAGGATGGCTCAATCTCGCTACTTCGATGCGCAGTCCGCGAA 5758

QY 28 Pro-----ProAsnArgCysLeuLysAspThr 36

Db 5757 TCTTGACCTGGAAGTCATTACGACCTAGATACTCGACCTCCCGCTGCTTCCAGGAATCC 5698

QY 37 AspCysProGlyIleLysLysCysCysGluGlySerCysGlyMetAlaCysPheVal 55

Db 5697 AGCGTACCAGGTACCGGCTCTGTGACAAACGGGAGTTACGACCTTCTCGCTTATCTT 5641

RESULT 27

US-08-222-617A-1/c

; Sequence 1, Application US/08222617A

; Patent No. 5882879

; GENERAL INFORMATION:

; APPLICANT: Veenstra, Annemarie E.

; APPLICANT: Martin, Juan F.

; APPLICANT: Garcia, Bruno D.

; APPLICANT: Gutierrez, Santiago

; APPLICANT: Barredo, Jose L.

; APPLICANT: Von Doehren, Hans

; APPLICANT: Palissa, Harriet

; APPLICANT: Van Liempt, Henk

; APPLICANT: Montenegro, Eduardo P.

; TITLE OF INVENTION: A Method for Influencing Beta-Lactam

; TITLE OF INVENTION: Antibiotic Production and for Isolation of Large

; TITLE OF INVENTION: Quantities of ACV Synthetase

; NUMBER OF SEQUENCES: 27

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: McDonnell Boehnen Hulbert & Berghoff

; STREET: 300 South Wacker Drive

; CITY: Chicago

; STATE: Illinois

; COUNTRY: USA

; ZIP: 60606

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: PatentIn Release #1.0, Version #1.25

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/222,617A

; FILING DATE: 04-APR-1994

; CLASSIFICATION: 435

; ATTORNEY/AGENT INFORMATION:

; REFERENCE/DOCKET NUMBER: 97,157

; INFORMATION FOR SEQ ID NO: 1:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 12364 base pairs

; TYPE: nucleic acid

; STRANDEDNESS: single

; TOPOLOGY: linear

; MOLECULE TYPE: DNA (genomic)

; HYPOTHETICAL: NO

; ANTI-SENSE: NO

; ORIGINAL SOURCE:

; ORGANISM: Penicillium chrysogenum

; FEATURE:

; NAME/KEY: CDS

; LOCATION: 264..11600

; OTHER INFORMATION: /function= "enzyme"

; OTHER INFORMATION: /product= "ACV Synthetase"

US-08-222-617A-1

Alignment Scores:

Pred. No.: 394 Length: 12364

Score: 66.50 Matches: 17

Percent Similarity: 35.44% Conservative: 11

Best Local Similarity: 21.52% Mismatches: 26

Query Match: 20.34% Indels: 25

DB: 2 Gaps: 2

09-833799-13B (1-57) x US-08-222-617A-1 (1-12364)

QY 2 GlnGluProValLysGlyProValSerThrLysProGlySerCysProIleIle----- 19

Db 5877 CAACCAGGAATTTCTGGGCCCCCTTCTCTGCAATCCTTGGCAATCACCACAGACTGTTTGA 5818
QY 20 -----LeulleArgCysAlaMetLeuAsn 27
|||:|||||:|:|:
Db 5817 TTCCGTGATAAGACGATAGGATGGCCTCAATCTCGCCTACTTCGATGGCAGTCCGGAA 5758
QY 28 Pro-----ProAsnArgCysLeuLysAspThr 36
|||:|||||:|:|:
Db 5757 TCTTGACCTGGAAGTCATTACGACCTAGATACTCGACCTCCCCGCTGCTTCCAGGAATCC 5698
QY 37 AspCysProGlyIleLysLysCysCysGlySerCysGlyMetAlaCysPheVal 55
|||:|||||:|:|:
Db 5697 AGCGTACCAGGTACCGGTCTTGTACAAACGGGAGTTACGACCTTCTCGCTTATCTT 5641
RESULT 28
US-07-963-538B-34/c
; Sequence 34, Application US/07963538B
; Patent No. 5851983
; GENERAL INFORMATION:
; APPLICANT: SUGIYAMA, TAKASHI
; APPLICANT: KAMIMURA, TAKASHI
; APPLICANT: MASUDA, KENICHI
; APPLICANT: OKADA, MASAHIRO
; APPLICANT: OHTSUKA, EIKO
; APPLICANT: IMAIZUMI, ATSUSHI
; APPLICANT: WATANABE, KUNIHITO
; APPLICANT: SUGA, TETSUYA
; APPLICANT: MATSUMOTO, YOHICHI
; APPLICANT: TAKEUCHI, AKIKO
; TITLE OF INVENTION: ELASTASE INHIBITORY POLYPEPTIDE AND
; TITLE OF INVENTION: PROCESS FOR PRODUCTION THEREOF BY RECOMBINANT GENE
; TITLE OF INVENTION: TECHNOLOGY
; NUMBER OF SEQUENCES: 36
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: COOLEY GODWARD LLP
; STREET: FIVE PALO ALTO SQUARE, 4TH FLOOR
; CITY: PALO ALTO
; STATE: CA
; COUNTRY: USA
; ZIP: 94306-2155
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/963,538B
; FILING DATE: 20-OCT-1992
; CLASSIFICATION: 530
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/843,359
; FILING DATE: 25-FEB-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/408,483
; FILING DATE: 22-AUG-1989
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JP 4-212399
; FILING DATE: 17-JUL-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JP 4-212398
; FILING DATE: 17-JUL-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JP 3-355553
; FILING DATE: 24-DEC-1991
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JP 62-330219
; FILING DATE: 28-DEC-1987
; ATTORNEY/AGENT INFORMATION:
; NAME: NEELEY PH.D., RICHARD L.
; REGISTRATION NUMBER: 30,092
; REFERENCE/DOCKET NUMBER: TEJN-005/02US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-843-5070

; TELEFAX: 415-857-0663
; TELEX: 380816 COOLEY PA
; INFORMATION FOR SEQ ID NO: 34:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 68 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single stranded
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (synthetic)
US-07-963-538B-34
Alignment Scores:
Pred. No.: 0.548 Length: 68
Score: 65.00 Matches: 12
Percent Similarity: 57.14% Conservative: 0
Best Local Similarity: 57.14% Mismatches: 9
Query Match: 19.88% Indels: 0
DB: 2 Gaps: 0
09-833799-13B (1-57) x US-07-963-538B-34 (1-68)
QY 25 MetLeuAsnProProAsnArgCysLeuLysAspThrAspCysProGlyIleLysLysCys 44
Db 66 ATGCTGAACCCGCCGAACCTTCTGTGAAATGGACGGTCAGTGTAAACGAGATCTGAAATGT 7
QY 45 Cys 45
Db 6 TGT 4
RESULT 29
US-08-306-691B-50/c
; Sequence 50, Application US/08306691B
; Patent No. 5734039
; GENERAL INFORMATION:
; APPLICANT: Calabretta, Bruno
; APPLICANT: Skorski, Tomasz
; TITLE OF INVENTION: ANTISENSE
; TITLE OF INVENTION: OLIGONUCLEOTIDES TARGETING COOPERATING ONCOGENES
; NUMBER OF SEQUENCES: 55
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Seidel, Gonda, Lavorgna & Monaco, P.C.
; STREET: Two Penn Center, Suite 1800
; CITY: Philadelphia
; STATE: Pennsylvania
; COUNTRY: U.S.A.
; ZIP: 19102
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.50 inch, 720 Kb
; COMPUTER: IBM PS/2
; OPERATING SYSTEM: MS-DOS
; SOFTWARE: WordPerfect 5.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/306,691B
; FILING DATE: September 15, 1994
; CLASSIFICATION: 514
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Monaco, Daniel A.
; REGISTRATION NUMBER: 30,480
; REFERENCE/DOCKET NUMBER: 8321-8
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (215) 568-8383
; TELEFAX: (215) 568-5549
; TELEX: No. 5734039e
; INFORMATION FOR SEQ ID NO: 50:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1364 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
US-08-306-691B-50

Alignment Scores:
Pred. No.: 36.4 Length: 1364
Score: 64.50 Matches: 17
Percent Similarity: 37.04% Conservative: 3
Best Local Similarity: 31.48% Mismatches: 15
Query Match: 19.72% Indels: 19
DB: 1 Gaps: 4

09-833799-13B (1-57) x US-08-306-691B-50 (1-1364)

QY 11 ThrLysProGlySerCysProIleLeuIleLeuArgCysAlaMetLeuAsnProProAsn 30
|||||
Db 800 ACAGAACTGGAAGGTGCCT-----TGTCCTTGT-----CCACCCACC 762

QY 31 ArgCysLeuLysAspThrAspCysProGlyIleLeuLysCysGluGlySerCysGly 50
|||||
Db 761 AGATGC-----TGTCCTTTCATGTCGCGCTGCGAGCAGGTCCAAACAGGA 717

QY 51 -----MetAlaCysPheValProGln 57
Db 716 ACTGGTACAGCGGATCTTCTTGTCTGCTGCTCCTCCCCAG 675

RESULT 30
PCT-US93-06251-65/c
; Sequence 65, Application PC/TUS9306251
; GENERAL INFORMATION:
; APPLICANT: Wickstrom, Eric and Rife, Jason P.
; TITLE OF INVENTION: Trivalent Synthesis of Oligonucleotides Containing
; TITLE OF INVENTION: Stereospecific Alkylphosphonates and Arylphosphonates
; NUMBER OF SEQUENCES: 93
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SCULLY, SCOTT, MURPHY & PRESSER
; STREET: 400 Garden City Plaza
; CITY: Garden City
; STATE: NY
; COUNTRY: USA
; ZIP: 11530
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US93/06251
; FILING DATE: 19930630
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: DiGiglio, Frank S.
; REGISTRATION NUMBER: 31,346
; REFERENCE/DOCKET NUMBER: 8586
; TELEPHONE: 516-742-4343
; TELEFAX: 516-742-4366
; TELEX: 230 901 SANS UR
; INFORMATION FOR SEQ ID NO: 65:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1364 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
PCT-US93-06251-65

Alignment Scores:
Pred. No.: 36.4 Length: 1364
Score: 64.50 Matches: 17
Percent Similarity: 37.04% Conservative: 3
Best Local Similarity: 31.48% Mismatches: 15
Query Match: 19.72% Indels: 19
DB: 5 Gaps: 4

09-833799-13B (1-57) x PCT-US93-06251-65 (1-1364)

QY 11 ThrLysProGlySerCysProIleLeuIleLeuArgCysAlaMetLeuAsnProProAsn 30
|||||
Db 800 ACAGAACTGGAAGGTGCCT-----TGTCCTTGT-----CCACCCACC 762

QY 31 ArgCysLeuLysAspThrAspCysProGlyIleLeuLysCysGluGlySerCysGly 50
|||||
Db 761 AGATGC-----TGTCCTTTCATGTCGCGCTGCGAGCAGGTCCAAACAGGA 717

QY 51 -----MetAlaCysPheValProGln 57
Db 716 ACTGGTACAGCGGATCTTCTTGTCTGCTGCTCCTCCCCAG 675

RESULT 31
US-08-961-527-184/c
; Sequence 184, Application US/08961527
; Patent No. 6420135
; GENERAL INFORMATION:
; APPLICANT: Charles Kunsch
; TITLE OF INVENTION: Streptococcus pneumoniae Polynucleotides and Sequenc
; NUMBER OF SEQUENCES: 391
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Human Genome Sciences, Inc.
; STREET: 9410 Key West Avenue
; CITY: Rockville
; STATE: Maryland
; COUNTRY: USA
; ZIP: 20850
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.50 inch, 1.4Mb storage
; COMPUTER: HP Vectra 486/33
; OPERATING SYSTEM: MSDOS version 6.2
; SOFTWARE: ASCII Text
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/961,527
; FILING DATE:
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Brookes, A. Anders
; REGISTRATION NUMBER: 36,373
; REFERENCE/DOCKET NUMBER: PB340P1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (301) 309-8504
; TELEFAX: (301) 309-8512
; INFORMATION FOR SEQ ID NO: 184:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1590 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
US-08-961-527-184

Alignment Scores:
Pred. No.: 44.7 Length: 1590
Score: 64.50 Matches: 16
Percent Similarity: 45.61% Conservative: 10
Best Local Similarity: 28.07% Mismatches: 29
Query Match: 19.72% Indels: 3
DB: 4 Gaps: 1

09-833799-13B (1-57) x US-08-961-527-184 (1-1590)

QY 1 AlaGlnGluProValLysGlyProValSerThrLysProGlySerCysProIleLeu 20
|||||
Db 780 GCTAAACGACCAATACGTTCTCCAATAGGAGTTGAACCTGTAAAGTTGATGAAGTTGACT 721

QY 21 IleArgCysAlaMetLeuAsnProProAsnArgCysLeuLysAspThrAspCysProGly 40
|||||
Db 720 TCTTTGTGCTCAATGATATA-ATCCCAAT-----TTCTGAACACGACCTGTAATGGT 668

100

1. The first step in the process is to identify the problem or issue that needs to be addressed. This involves gathering information and understanding the context of the problem.

Db 1252 TCGTGTGCGATTGAGGTAGTCGCATAGTGATCCGTTATGCTGGTAGGTGATATCAGCC 1193
QY 27 AsnProProAsnArgCysLeuLysAspThrAspCysProGlyIleLysLysCysCysGlu 46
Db 1192 AATATTCCG---GCTTGTCATGTGCTTCTCAACGCCAGGAATTCGAGGATGTTCCGGAT 1136
QY 47 GlySerCysGlyMetAlaCys 53
Db 1135 GCGCATGCGCGCAGCTTGT 1115
RESULT 34
US-08-459-951-1/c
; Sequence 1, Application US/08459951
; Patent No. 6093547
; GENERAL INFORMATION:
; APPLICANT: JIN, DONALD F
; APPLICANT: OPPERMANN, HERMANN
; APPLICANT: KUBERASAMPATH, THANGAVEL
; APPLICANT: SMART, JOHN E
; TITLE OF INVENTION: NOVEL MORPHOGEN CELL SURFACE RECEPTOR
; NUMBER OF SEQUENCES: 12
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: PATENT ADMINISTRATOR, CREATIVE BIOMOLECULES,
; ADDRESSEE: INC
; STREET: 45 SOUTH STREET
; CITY: HOPKINTON
; STATE: MA
; COUNTRY: USA
; ZIP: 01748
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/459,951
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/357,533
; FILING DATE: 16-DEC-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: KELLY, ROBIN D
; REGISTRATION NUMBER: 34,637
; REFERENCE/DOCKET NUMBER: CRP-073FW
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (508)-435-9001
; TELEFAX: (508)-435-0992
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2625 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 379..1929
; OTHER INFORMATION: /product= "DROSOPHILA MORPHOGEN
; OTHER INFORMATION: RECEPTOR"
US-08-459-951-1

Alignment Scores:
Pred. No.: 88 Length: 2625
Score: 64.50 Matches: 23
Percent Similarity: 44.78% Conservative: 7
Best Local Similarity: 34.33% Mismatches: 22
Query Match: 19.72% Indels: 15
DB: 3 Gaps: 4
09-833799-13B (1-57) x US-08-459-951-1 (1-2625)

QY 1 AlaGlnGluProValLysGlyProValSerThrLysProGly----- 14
Db 1312 GCAGATGTGCCAGTCCATTGGCCATGGACTGAGCTCAGCGATCGGCACAACTCTGGCCATGAGA 1253
QY 15 SerCysProIle-----IleLeulleArgCysAla-----MetLeu 26
Db 1252 TCGTGTGCGATTGAGGTAGTCGCATAGTATCCGTTATGCTGGTAGGTGATATCAGCC 1193
QY 27 AsnProProAsnArgCysLeuLysAspThrAspCysProGlyIleLysLysCysCysGlu 46
Db 1192 AATATTCCG---GCTTGTCATGTGCTTCTCAACGCCAGGAATTCGAGGATGTTCCGGAT 1136
QY 47 GlySerCysGlyMetAlaCys 53
Db 1135 GCGCATGCGCGCAGCTTGT 1115
RESULT 35
US-09-036-987A-1/c
; Sequence 1, Application US/09036987A
; Patent No. 6143526
; GENERAL INFORMATION:
; APPLICANT: Baltz, Richard H.
; APPLICANT: Broughton, Mary C.
; APPLICANT: Crawford, Kathryn P.
; APPLICANT: Madduri, Krishnamurthy
; APPLICANT: Merlo, Donald J.
; APPLICANT: Treadway, Patti J.
; APPLICANT: Turner, Jan R.
; APPLICANT: Waldron, Clive
; TITLE OF INVENTION: Biosynthetic Genes For Spinosyn Insecticide
; NUMBER OF SEQUENCES: 39
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Dow AgroSciences LLC Patent Department
; STREET: 9330 Zionsville Road
; CITY: Indianapolis
; STATE: Indiana
; COUNTRY: USA
; ZIP: 46268
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/036,987A
; FILING DATE: 09-MAR-1998
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Stuart, Donald R
; REGISTRATION NUMBER: 28,479
; REFERENCE/DOCKET NUMBER: 50,608
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (317)337-4816
; TELEFAX: (317)337-4847
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 80161 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
US-09-036-987A-1

Alignment Scores:
Pred. No.: 1.02e+04 Length: 80161
Score: 64.00 Matches: 20
Percent Similarity: 41.94% Conservative: 6
Best Local Similarity: 32.26% Mismatches: 17
Query Match: 19.57% Indels: 20
DB: 3 Gaps: 4
09-833799-13B (1-57) x US-09-036-987A-1 (1-80161)

APPLICANT: Young, Michael W.
 APPLICANT: Sehgal, Amita
 APPLICANT: Voshall, Leslie B.
 APPLICANT: Price, Jeffrey L.
 APPLICANT: Myers, Michael
 TITLE OF INVENTION: NUCLEAR LOCALIZATION FACTOR ASSOCIATED
 TITLE OF INVENTION: WITH CIRCADIAN RHYTHMS
 NUMBER OF SEQUENCES: 8
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: Klauber & Jackson
 STREET: 411 Hackensack Avenue
 CITY: Hackensack
 STATE: New Jersey
 COUNTRY: USA
 ZIP: 07601
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: PatentIn Release #1.0, Version #1.25
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/619,198
 FILING DATE: 20-MAR-1996
 CLASSIFICATION: 514
 ATTORNEY/AGENT INFORMATION:
 NAME: Jackson Esq., David A.
 REGISTRATION NUMBER: 26,742
 REFERENCE/DOCKET NUMBER: 600-1-128A CP1
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: 201 487-5800
 TELEFAX: 201 343-1684
 TELEX: 133521
 INFORMATION FOR SEQ ID NO: 1:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 5198 base pairs
 TYPE: nucleic acid
 STRANDEDNESS: double
 TOPOLOGY: linear
 MOLECULE TYPE: cDNA
 DESCRIPTION: tim cDNA sequence wherein R at position 157
 HYPOTHETICAL: NO
 ANTI-SENSE: NO
 ORIGINAL SOURCE:
 ORGANISM: Drosophila melanogaster
 US-08-619-198-1

```

Alignment Scores:
Pred. NO.: 297
Score: 63.50
Percent Similarity: 41.86%
Best Local Similarity: 32.56%
Query Match: 19.42%
DB: 2

09-833799-13B (1-57) x US-08-619-198-1 (1-5198)

QY 22 ArgCysAlaMetLeuAsnProProAsnArgCysLeuLysAspThrAsp----- 37
      ||||| ::||| ||||| ||||| ::|
Db 4396 AGGTGTACCTCCTCCGCCTGCCGCGCGGTGCTCATCAGAGCCACCATGGACAGGGAGG 4337
      ||||| ::||| ||||| ||||| |||||
QY 38 -----CysProGlyIleLysLysCysCysGluGlySerCysGly----- 50
      ||||| ::||| ||||| ||||| |||||
Db 4336 AGACGGTGCCCCCGCAACCAGAACTGCTGCTGTTCTCGTTGGCTTCCCATGAAGGTGT 4277
      ||||| ::||| ||||| ||||| |||||
QY 51 MetAlaCys 53
      ::|||
Db 4276 TTAGCATGT 4268

RESULT 40
US-09-028-934-28
; Sequence 28, Application US/09028934
; Patent No. 6117670
; GENERAL INFORMATION:

```

RESULT 39
US-08-619-198-1/c
; Sequence 1, Application US/08619198
; Patent No. 5885831
; GENERAL INFORMATION:

Db 1407 TGCTCCGCCAGCACAGAGCCCTGGC-----AGGTGTACGGCTTCGTGCGGGCCT 1457

Qy 52 AlaCys-----PheValProGln 57
|||||:::|||||

Db 1458 GCCTGCGCGCGGTGCTGCTGCCAG 1481

RESULT 43

US-08-974-549A-1

; Sequence 1, Application US/08974549A

; Patent No. 6166178

; GENERAL INFORMATION:

; APPLICANT: Cech, Thomas R.

; APPLICANT: Lingner, Joachim

; APPLICANT: Nakamura, Toru

; APPLICANT: Chapman, Karen B.

; APPLICANT: Morin, Gregg B.

; APPLICANT: Harley, Calvin B.

; APPLICANT: Andrews, William H.

; TITLE OF INVENTION: Human Telomerase Catalytic Subunit

; NUMBER OF SEQUENCES: 727

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Townsend and Townsend and Crew LLP

; STREET: Two Embarcadero Center, Eighth Floor

; CITY: San Francisco

; STATE: California

; COUNTRY: USA

; ZIP: 94111-3834

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: PatentIn Release #1.0, Version #1.30

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/974,549A

; FILING DATE: 19-NOV-1997

; CLASSIFICATION: 536

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: US 08/724,643

; FILING DATE: 01-OCT-1996

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: US 08/844,419

; FILING DATE: 18-APR-1997

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: US 08/846,017

; FILING DATE: 25-APR-1997

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: US 08/851,843

; FILING DATE: 06-MAY-1997

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: US 08/854,050

; FILING DATE: 09-MAY-1997

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: US 08/911,312

; FILING DATE: 14-AUG-1997

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: US 08/912,951

; FILING DATE: 14-AUG-1997

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: US 08/915,503

; FILING DATE: 14-AUG-1997

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: WO PCT/US97/17618

; FILING DATE: 01-OCT-1997

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: WO PCT/US97/17885

; FILING DATE: 01-OCT-1997

; ATTORNEY/AGENT INFORMATION:

; NAME: Apple, Randolph Ted

; REGISTRATION NUMBER: 36,429

; REFERENCE/DOCKET NUMBER: 015389-002610US

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: (415) 576-0200

; TELEFAX: (415) 576-0300

; INFORMATION FOR SEQ ID NO: 1:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 4015 base pairs

; TYPE: nucleic acid

; STRANDEDNESS: single

; TOPOLOGY: linear

; MOLECULE TYPE: cDNA

; FEATURE:

; NAME/KEY: CDS

; LOCATION: 56..3454

; OTHER INFORMATION: /product= "hTRT"

; OTHER INFORMATION: /note= "human telomerase reverse

; OTHER INFORMATION: transcriptase (hTRT) catalytic protein

; OTHER INFORMATION: component"

US-08-974-549A-1

Alignment Scores:

Pred. No.: 282 Length: 4015

Score: 62.50 Matches: 23

Percent Similarity: 38.24% Conservative: 3

Best Local Similarity: 33.82% Mismatches: 17

Query Match: 19.11% Indels: 25

DB: 4 Gaps: 5

09-833799-13B (1-57) x US-08-974-549A-1 (1-4015)

Qy 2 GlnGluProValLysGlyProValSerThrLysProGlySerCysProIleLeuIle 21
|||:::|||||:::|||||

Db 1317 CAGCAGCGCGTG-----TCTGTGCCCGGAGAGAGCCCGAGGCTCTGTGG 1361

Qy 22 ArgCysAlaMetLeuAsnProProAsnArg----- 31

Db 1362 CGG-----CCCCCGAGGAGGAGACACAGACCCCGCTCGCTGGTGCAGC 1406

Qy 32 CysLeuLysAspThrAspCysProGlyIleLysLysCysCysGluGlySerCysGlyMet 51
|||:::|||||:::|||||

Db 1407 TGCTCCGCCAGCACAGAGCCCTGGC-----AGGTGTACGGCTTCGTGCGGGCCT 1457

Qy 52 AlaCys-----PheValProGln 57
|||||:::|||||

Db 1458 GCCTGCGCGCGGTGCTGCTGCCAG 1481

RESULT 44

US-08-854-050-224

; Sequence 224, Application US/08854050

; Patent No. 6261836

; GENERAL INFORMATION:

; APPLICANT: Cech, Thomas R.

; APPLICANT: Lingner, Joachim

; APPLICANT: Nakamura, Toru

; APPLICANT: Chapman, Karen B.

; APPLICANT: Morin, Gregg B.

; APPLICANT: Harley, Calvin

; APPLICANT: Andrews, William H.

; TITLE OF INVENTION: No. 6261836el Telomerase

; NUMBER OF SEQUENCES: 225

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Townsend and Townsend and Crew LLP

; STREET: Two Embarcadero Center, 8th Floor

; CITY: San Francisco

; STATE: California

; COUNTRY: United States of America

; ZIP: 94111

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: PatentIn Release #1.0, Version #1.30

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/854,050

; FILING DATE: 09-MAY-1997

; CLASSIFICATION: 536

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: US 08/724,643

; FILING DATE: 01-OCT-1996

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: US 08/844,419

; FILING DATE: 18-APR-1997

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: US 08/846,017

; FILING DATE: 25-APR-1997

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: US 08/851,843

; FILING DATE: 06-MAY-1997

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: US 08/854,050

; FILING DATE: 09-MAY-1997

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: US 08/911,312

; FILING DATE: 14-AUG-1997

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: US 08/912,951

; FILING DATE: 14-AUG-1997

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: US 08/915,503

; FILING DATE: 14-AUG-1997

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: WO PCT/US97/17618

; FILING DATE: 01-OCT-1997

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: WO PCT/US97/17885

; FILING DATE: 01-OCT-1997

; ATTORNEY/AGENT INFORMATION:

; NAME: Apple, Randolph Ted

; REGISTRATION NUMBER: 36,429

; REFERENCE/DOCKET NUMBER: 015389-002610US

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: (415) 576-0200

; TELEFAX: (415) 576-0300

RESULT 47
US-09-128-354-1
; Sequence 1, Application US/09128354
; Patent No. 6337200
; GENERAL INFORMATION:
; APPLICANT: Morin, Gregg B.
; APPLICANT: Geron Corporation
; TITLE OF INVENTION: Human Telomerase Catalytic Subunit Variants
; FILE REFERENCE: 015389-00310US
; CURRENT APPLICATION NUMBER: US/09/128,354
; CURRENT FILING DATE: 1998-08-03
; EARLIER APPLICATION NUMBER: US 08/851,843
; EARLIER FILING DATE: 1997-05-06
; EARLIER APPLICATION NUMBER: US 08/854,050
; EARLIER FILING DATE: 1997-05-09
; EARLIER APPLICATION NUMBER: US 08/911,312

LENGTH: 4015

```

; SEQ ID NO 1
: LENGTH: 4015

```

; SEQ ID

EARLIER


```

; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 56..3454
; OTHER INFORMATION: /note= "refined sequence of hTERT cDNA"
US-08-974-549A-343

Alignment Scores:
Pred. No.: 284 Length: 4037
Score: 62.50 Matches: 23
Percent Similarity: 38.24% Conservative: 3
Best Local Similarity: 33.82% Mismatches: 17
Query Match: 19.11% Indels: 25
DB: 4 Gaps: 5

09-833799-13B (1-57) x US-08-974-549A-343 (1-4037)

QY 2 GlnGluProValLysGlyProValSerThrLysProGlySerCysProIleIleLeuIle 21
|||:::||||| ::: ||||| |||
Db 1317 CAGCAGCCGGTG-----TCTGTCCCGGGAGAGAGCCCGAGGGCTCTGTGG 1361

QY 22 ArgCysAlaMetLeuAsnProProAsnArg----- 31
||| ||||| |||
Db 1362 CGG-----CCCCGAGGAGGAGACAGACCCCGTCGCTGGTGCAGC 1406

QY 32 CysLeuLysAspThrAspCysProGlyIleLysLysCysCysGluGlySerCysGlyMet 51
||| ||| ||| ||||| |||
Db 1407 TGCTCCGCCAGCACAGCAGCCCTGGC-----AGGTGTACGGCTTCGTGCGGGCCT 1457

QY 52 AlaCys-----PheValProGln 57
||||| ::: |||||
Db 1458 GCCTGCGCGGCTGGTGCCTCCAG 1481

RESULT 51
US-08-974-549A-6
; Sequence 6, Application US/08974549A
; Patent No. 6166178
; GENERAL INFORMATION:
; APPLICANT: Cech, Thomas R.
; APPLICANT: Lingner, Joachim
; APPLICANT: Nakamura, Toru
; APPLICANT: Chapman, Karen B.
; APPLICANT: Morin, Gregg B.
; APPLICANT: Harley, Calvin B.
; APPLICANT: Andrews, William H.
; TITLE OF INVENTION: Human Telomerase Catalytic Subunit
; NUMBER OF SEQUENCES: 727
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Townsend and Townsend and Crew LLP
; STREET: Two Embarcadero Center, Eighth Floor
; CITY: San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94111-3834
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/974,549A
; FILING DATE: 19-NOV-1997
; CLASSIFICATION: 536
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/724,643
; FILING DATE: 01-OCT-1996
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/844,419
; FILING DATE: 18-APR-1997
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/846,017

```


; Sequence 6, Application US/09467997
; Patent No. 6379925
; GENERAL INFORMATION:
; APPLICANT: Kitajewski, Jan
; APPLICANT: Uyttendaele, Hendrik
; TITLE OF INVENTION: ANGIOGENIC MODULATION BY NOTCH SIGNAL TRANSDUCTION
; FILE REFERENCE: 53863-A-PCT-US
; CURRENT APPLICATION NUMBER: US/09/467,997
; CURRENT FILING DATE: 1999-12-20
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 6
; LENGTH: 6677
; TYPE: DNA
; ORGANISM: mouse
US-09-467-997-6

Alignment Scores:
Pred. No.: 561 Length: 6677
Score: 62.50 Matches: 20
Percent Similarity: 40.30% Conservative: 7
Best Local Similarity: 29.85% Mismatches: 11
Query Match: 19.11% Indels: 29
DB: 4 Gaps: 4

09-833799-13B (1-57) x US-09-467-997-6 (1-6677)

QY 10 SerThrLysProGlySerCysProIleIleLeuIleArgCysAlaMetLeuAsn----- 27
Db 3647 TCCCTAAGCCAGGTCACACCGCTC-----TGTCCTGCCTCAGTGGTTT 3694
QY 28 -----ProProAsnArgCys 32
Db 3695 GGGGGCCCTGACTGTCTGACACCTCCAGCTCCACCGGGCTGGGTCCCGCTCACCTGC 3754
QY 33 LeuLysAsp-----ThrAspCysProGlyIleLysLysCysGluGlySerCys 49
Db 3755 CTGCACAATGGTACCTGCACTGAGACCCCTGGGTG-----GGCAACCCG 3799
QY 50 GlyMetAlaCysPheValPro 56
Db 3800 GGCTTTCAATGCACCTGCCCT 3820

RESULT 54
US-08-400-159-3/c
; Sequence 3, Application US/08400159
; Patent No. 5869282
; GENERAL INFORMATION:
; APPLICANT: Ish-Horowicz, David
; APPLICANT: Henrique, Domingos M.P.
; APPLICANT: Lewis, Julian H.
; APPLICANT: Myat, Anna M.
; APPLICANT: Fleming, Robert J.
; APPLICANT: Artavanis-Tsakonas, Spyridon
; APPLICANT: Mann, Robert S.
; APPLICANT: Gray, Grace E.
; TITLE OF INVENTION: NUCLEOTIDE AND PROTEIN SEQUENCES OF THE
; TITLE OF INVENTION: SERRATE GENE AND METHODS BASED THEREON
; NUMBER OF SEQUENCES: 20
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Pennie & Edmonds
; STREET: 1155 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: USA
; ZIP: 10036-2711
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/400,159

; FILING DATE: 07-MAR-1995
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Misrock, S. Leslie
; REGISTRATION NUMBER: 18,872
; REFERENCE/DOCKET NUMBER: 7326-029
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 790-9090
; TELEFAX: (212) 869-9741/8864
; TELEX: 66141 PENNIE
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 708 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: unknown
; MOLECULE TYPE: cDNA
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 1..708
US-08-400-159-3
Alignment Scores:
Pred. No.: 36.4 Length: 708
Score: 61.50 Matches: 12
Percent Similarity: 43.75% Conservative: 2
Best Local Similarity: 37.50% Mismatches: 11
Query Match: 18.81% Indels: 7
DB: 2 Gaps: 1
09-833799-13B (1-57) x US-08-400-159-3 (1-708)
QY 13 ProGlySerCysProIleIleLeuIleArgCysAlaMetLeuAsnProProAsnArgCys 32
Db 175 CCAGGCACTTGCCG-----TCGCTCCGTCGCGACTCCCGCTGC 137
QY 33 LeuLysAspThrAspCysProGlyIleLysLysCys 44
Db 136 AGCAGCGACCTCGTGTCTCCCGCCCGTGATCGTTGC 101

RESULT 55
US-08-611-729A-3/c
; Sequence 3, Application US/08611729A
; Patent No. 6004924
; GENERAL INFORMATION:
; APPLICANT: Ish-Horowicz, David
; APPLICANT: Henrique, Domingos M.P.
; APPLICANT: Lewis, Julian H.
; APPLICANT: Myat, Anna M.
; APPLICANT: Fleming, Robert J.
; APPLICANT: Artavanis-Tsakonas, Spyridon
; APPLICANT: Mann, Robert S.
; APPLICANT: Gray, Grace E.
; TITLE OF INVENTION: NUCLEOTIDE AND PROTEIN SEQUENCES OF THE
; TITLE OF INVENTION: SERRATE GENE AND METHODS BASED THEREON
; NUMBER OF SEQUENCES: 20
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Pennie & Edmonds
; STREET: 1155 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: U.S.A.
; ZIP: 10036-2711
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/611,729A
; FILING DATE: 06-MAR-1996
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:

; NAME: Misrock, S. Leslie
; REGISTRATION NUMBER: 18,872
; REFERENCE/DOCKET NUMBER: 7326-037
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 790-9090
; TELEFAX: (212) 869-9741/8864
; TELEX: 66141 PENNIE
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 708 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: unknown
; MOLECULE TYPE: cDNA
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 1..708
; US-08-611-729A-3

Alignment Scores:
Pred. No.: 36.4 Length: 708
Score: 61.50 Matches: 12
Percent Similarity: 43.75% Conservative: 2
Best Local Similarity: 37.50% Mismatches: 11
Query Match: 18.81% Indels: 7
DB: 3 Gaps: 1

09-833799-13B (1-57) x US-08-611-729A-3 (1-708)

QY 13 ProGlySerCysProIleLeuIleArgCysAlaMetLeuAsnProProAsnArgCys 32
|||||:|||||
Db 175 CCAGGCACTTGCCCG-----TCGCTCCGTCCGACTCCCGCTGC 137
QY 33 LeuLysAspThrAspCysProGlyIleLysLysCys 44
|||:|||||
Db 136 AGCAGCGACCTCGTTGTCCCGCCCGCGTGATCGTTGC 101

RESULT 56
US-08-886-152-2/c
; Sequence 2, Application US/08886152
; Patent No. 5880273
; GENERAL INFORMATION:
; APPLICANT: ADACHI, HIDEKI
; APPLICANT: TSUJIMOTO, MASAFUMI
; APPLICANT: ARAI, HIROYUKI
; APPLICANT: INOUE, KEIZO
; TITLE OF INVENTION: PLATELET ACTIVATING FACTOR
; TITLE OF INVENTION: ACETYLHYDROLASE, AND GENE THEREOF
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: OBLON, SPIVAK, MCCLELLAND, MAIER & NEUSTADT,
; ADDRESSEE: P.C.
; STREET: 1755 S. JEFFERSON DAVIS HIGHWAY, SUITE 400
; CITY: ARLINGTON
; STATE: VA
; COUNTRY: USA
; ZIP: 22202
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/886,152
; FILING DATE: 30-JUN-1997
; CLASSIFICATION: 536
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JP 8-188369
; FILING DATE: 28-JUN-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: OBLON, NORMAN F.
; REGISTRATION NUMBER: 24,618
; REFERENCE/DOCKET NUMBER: 2292-041-0

; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 703-413-3000
; TELEFAX: 703-413-2220
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1665 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; ORIGINAL SOURCE:
; ORGANISM: BOVINE (Bos taurus)
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 111..1286
; US-08-886-152-2

Alignment Scores:
Pred. No.: 116 Length: 1665
Score: 61.50 Matches: 13
Percent Similarity: 42.50% Conservative: 4
Best Local Similarity: 32.50% Mismatches: 14
Query Match: 18.81% Indels: 9
DB: 2 Gaps: 2

09-833799-13B (1-57) x US-08-886-152-2 (1-1665)

QY 13 ProGlySerCysProIleLeuIleArgCysAlaMetLeu----- 26
|||||:|||||
Db 823 CCAATGAATGTCCCATACAGCCACACGGCTCAGTCAATGCCGCCCTTCAAGTTCATC 764
QY 27 -----AsnProProAsnArgCysLeuLysAspThrAspCysProGlyIleLysLysCys 44
|||||:|||||
Db 763 AGATCCAATCCGCCAGGCAAGATGTTGAGA---ACGGCTGCCACGACGTGACCTCTTGT 707

RESULT 57
US-09-196-222-2/c
; Sequence 2, Application US/09196222
; Patent No. 6323017
; GENERAL INFORMATION:
; APPLICANT: ADACHI, HIDEKI
; APPLICANT: TSUJIMOTO, MASAFUMI
; APPLICANT: ARAI, HIROYUKI
; APPLICANT: INOUE, KEIZO
; TITLE OF INVENTION: PLATELET ACTIVATING FACTOR
; TITLE OF INVENTION: ACETYLHYDROLASE, AND GENE THEREOF
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: OBLON, SPIVAK, MCCLELLAND, MAIER & NEUSTADT,
; ADDRESSEE: P.C.
; STREET: 1755 S. JEFFERSON DAVIS HIGHWAY, SUITE 400
; CITY: ARLINGTON
; STATE: VA
; COUNTRY: USA
; ZIP: 22202
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/196,222
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/886,152
; FILING DATE: 30-JUN-1997
; APPLICATION NUMBER: JP 8-188369
; FILING DATE: 28-JUN-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: OBLON, NORMAN F.
; REGISTRATION NUMBER: 24,618
; REFERENCE/DOCKET NUMBER: 2292-041-0

TELECOMMUNICATION INFORMATION:
TELEPHONE: 703-413-3000
TELEFAX: 703-413-2220
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 1665 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA
ORIGINAL SOURCE:
ORGANISM: BOVINE (Bos taurus)
FEATURE:
NAME/KEY: CDS
LOCATION: 111..1286
US-09-196-222-2

Alignment Scores:
Pred. No.: 116 Length: 1665
Score: 61.50 Matches: 13
Percent Similarity: 42.50% Conservative: 4
Best Local Similarity: 32.50% Mismatches: 14
Query Match: 18.81% Indels: 9
DB: 4 Gaps: 2

09-833799-13B (1-57) x US-09-196-222-2 (1-1665)

QY 13 ProGlySerCysProIleIleLeuIleArgCysAlaMetLeu----- 26
||| ||||| ||| ::
Db 823 CCAATGAATGTCCATTACAGCCACACGGCTCAGTCAATGCCGCCCTTCAAGGTCATC 764
QY 27 -----AsnProAsnArgCysLeuLysAspThrAspCysProGlyIleLysLysCys 44
||| ||||| ||| ::
Db 763 AGATCCAATCCGCCAGGCAAGATGTTGAGA--ACGCGCTGCCCGCAGTGACCTCTTGT 707

RESULT 58

US-08-264-534-5/c
Sequence 5, Application US/08264534
Patent No. 5648464
GENERAL INFORMATION:
APPLICANT: Artavanis-Tsakonas, Spyridon et al.
TITLE OF INVENTION: Human No. 5648464ch And Delta, Binding Domains
TITLE OF INVENTION: In Toporythmic Proteins, And Methods Based Thereon
NUMBER OF SEQUENCES: 34
CORRESPONDENCE ADDRESS:
ADDRESSEE: Pennie & Edmonds
STREET: 1155 Avenue of the Americas
CITY: New York
STATE: New York
COUNTRY: U.S.A.
ZIP: 10036
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/264,534
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/695,189
FILING DATE: 03-MAY-1991
ATTORNEY/AGENT INFORMATION:
NAME: Misrock, S. Leslie
REGISTRATION NUMBER: 18,872
REFERENCE/DOCKET NUMBER: 7326-004
TELECOMMUNICATION INFORMATION:
TELEPHONE: 212 790-9090
TELEFAX: 212 8698864/9741
TELEX: 66141 PENNIE
INFORMATION FOR SEQ ID NO: 5:
SEQUENCE CHARACTERISTICS:

LENGTH: 2892 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: unknown
MOLECULE TYPE: cDNA
FEATURE:
NAME/KEY: CDS
LOCATION: 142..2640
US-08-264-534-5

Alignment Scores:
Pred. No.: 244 Length: 2892
Score: 61.50 Matches: 12
Percent Similarity: 43.75% Conservative: 2
Best Local Similarity: 37.50% Mismatches: 11
Query Match: 18.81% Indels: 7
DB: 1 Gaps: 1

09-833799-13B (1-57) x US-08-264-534-5 (1-2892)

QY 13 ProGlySerCysProIleIleLeuIleArgCysAlaMetLeuAsnProProAsnArgCys 32
||| ||||| ||| ::
Db 316 CCAGGCACTTGCCCG-----TCGCTCCGTCGCGACTCCCGCGTGC 278
QY 33 LeuLysAspThrAspCysProGlyIleLysLysCys 44
||| ||||| ||| ::
Db 277 AGCAGCGACCTCGTTGTCCCGCCCGTGATCGTTGC 242

RESULT 59

US-08-083-590A-1/c
Sequence 1, Application US/08083590A
Patent No. 5786158
GENERAL INFORMATION:
APPLICANT: Artavanis-Tsakonas, S. et al.
TITLE OF INVENTION: Therapeutic And Diagnostic Methods
TITLE OF INVENTION: And Compositions Based On No. 5786158ch Proteins And
TITLE OF INVENTION: Nucleic Acids
NUMBER OF SEQUENCES: 21
CORRESPONDENCE ADDRESS:
ADDRESSEE: Pennie & Edmonds
STREET: 1155 Avenue of the Americas
CITY: New York
STATE: New York
COUNTRY: U.S.A.
ZIP: 10036
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/083,590A
FILING DATE: 25-JUN-1993
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Misrock, S. Leslie
REGISTRATION NUMBER: 18,872
REFERENCE/DOCKET NUMBER: 7326-015
TELECOMMUNICATION INFORMATION:
TELEPHONE: 212 790-9090
TELEFAX: 212 8698864/9741
TELEX: 66141 PENNIE
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 2892 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: unknown
MOLECULE TYPE: cDNA
FEATURE:
NAME/KEY: CDS
LOCATION: 142..2640
US-08-083-590A-1

Alignment Scores: 244 Length: 2892
Pred. No.: 61.50 Matches: 12
Score: 43.75% Conservative: 2
Percent Similarity: 37.50% Mismatches: 11
Best Local Similarity: 18.81% Indels: 7
Query Match: 1 Gaps: 1
DB:

09-833799-13B (1-57) x US-08-083-590A-1 (1-2892)

QY 13 ProGlySerCysProIleIleLeuIleArgCysAlaMetLeuAsnProProAsnArgCys 32
|||||:|||||
Db 316 CCAGGCACTTGCCG-----TCGCTCCGTCGACTCCCGCTGC 278

QY 33 LeuLysAspThrAspCysProGlyIleLysLysCys 44
|||||:|||||
Db 277 AGCAGCACCTCGTTGTCGCCGCCGTCGATCGTTGC 242

RESULT 60
US-08-465-500-5/c
; Sequence 5, Application US/08465500
; Patent No. 5789195
; GENERAL INFORMATION:
; APPLICANT: Artavanis-Tsakonas, Spyridon
; APPLICANT: Muskavitch, Marc A.T.
; APPLICANT: Fehon, Richard G.
; APPLICANT: Rebay, Ilaria
; APPLICANT: Blaumueller, Cristine M.
; APPLICANT: Shepard, Scott B.
; TITLE OF INVENTION: HUMAN NOTCH AND DELTA, BINDING DOMAINS
; TITLE OF INVENTION: IN TOPOTHYMIC PROTEINS, AND METHODS BASED THEREON
; NUMBER OF SEQUENCES: 34
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: PENNIE & EDMONDS
; STREET: 1155 Avenue of the Americas
; CITY: New York
; STATE: NY
; COUNTRY: USA
; ZIP: 10036-2711
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/465,500
; FILING DATE: 05-JUN-1995
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Misrock, S. Leslie
; REGISTRATION NUMBER: 18,872
; REFERENCE/DOCKET NUMBER: 7326-034
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 790-9090
; TELEFAX: (212) 869-8864/9741
; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2892 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: unknown
; MOLECULE TYPE: CDNA
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 142..2640
US-08-465-500-5

Alignment Scores: 244 Length: 2892
Pred. No.: 61.50 Matches: 12
Score: 43.75% Conservative: 2
Percent Similarity: 37.50% Mismatches: 11
Best Local Similarity: 37.50%

Query Match: 18.81% Indels: 7
DB: 1 Gaps: 1

09-833799-13B (1-57) x US-08-465-500-5 (1-2892)

QY 13 ProGlySerCysProIleIleLeuIleArgCysAlaMetLeuAsnProProAsnArgCys 32
|||||:|||||
Db 316 CCAGGCACTTGCCG-----TCGCTCCGTCGACTCCCGCTGC 278

QY 33 LeuLysAspThrAspCysProGlyIleLysLysCys 44
|||||:|||||
Db 277 AGCAGCACCTCGTTGTCGCCGCCGTCGATCGTTGC 242

RESULT 61
US-08-346-126-5/c
; Sequence 5, Application US/08346126
; Patent No. 5849869
; GENERAL INFORMATION:
; APPLICANT: Artavanis-Tsakonas, Spyridon et al.
; TITLE OF INVENTION: Human No. 5849869ch And Delta, Binding Domains
; TITLE OF INVENTION: In Toporythmic Proteins, And Methods Based Thereon
; NUMBER OF SEQUENCES: 30
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Pennie & Edmonds
; STREET: 1155 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: U.S.A.
; ZIP: 10036
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/346,126
; FILING DATE:
; CLASSIFICATION: 530
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/791,923
; FILING DATE: 14-NOV-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: Misrock, S. Leslie
; REGISTRATION NUMBER: 18,872
; REFERENCE/DOCKET NUMBER: 7326-007
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212 790-9090
; TELEFAX: 212 8698864/9741
; TELEX: 66141 PENNIE
; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2892 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: unknown
; MOLECULE TYPE: CDNA
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 142..2640
US-08-346-126-5

Alignment Scores: 244 Length: 2892
Pred. No.: 61.50 Matches: 12
Score: 43.75% Conservative: 2
Best Local Similarity: 37.50% Mismatches: 11
Query Match: 18.81% Indels: 7
DB: 2 Gaps: 1

09-833799-13B (1-57) x US-08-346-126-5 (1-2892)

QY 13 ProGlySerCysProIleIleLeuIleArgCysAlaMetLeuAsnProProAsnArgCys 32
|||||:|||||

Db 316 CCAGGCACTTGCCCG-----TCGCTCCGTCCGACTCCCGGCTGC 278

QY 33 LeuLysAspThrAspCysProGlyIleLysLysCys 44
|||
Db 277 AGCAGCGACCTCGTTGTCCCGCCCGTGATCGTTGC 242

RESULT 62
US-08-346-128-5/c
; Sequence 5, Application US/08346128
; Patent No. 5856441
; GENERAL INFORMATION:
; APPLICANT: Artavanis-Tsakonas, Spyridon et al.
; TITLE OF INVENTION: Human No. 5856441ch And Delta, Binding Domains
; TITLE OF INVENTION: In Toporythmic Proteins, And Methods Based Thereon
; NUMBER OF SEQUENCES: 37
; CORRESPONDENCE ADDRESSES:
; ADDRESSEE: Pennie & Edmonds
; STREET: 1155 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: U.S.A.
; ZIP: 10036
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/346,128
; FILING DATE:
; CLASSIFICATION: 530
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/879,038
; FILING DATE: 30-APR-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: Misrock, S. Leslie
; REGISTRATION NUMBER: 18,872
; REFERENCE/DOCKET NUMBER: 7326-009
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212 790-9090
; TELEFAX: 212 8698864/9741
; TELEX: 66141 PENNIE
; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2892 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: unknown
; MOLECULE TYPE: cDNA
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 142..2640
US-08-346-128-5

Alignment Scores:
Pred. No.: 244 Length: 2892
Score: 61.50 Matches: 12
Percent Similarity: 43.75% Conservative: 2
Best Local Similarity: 37.50% Mismatches: 11
Query Match: 18.81% Indels: 7
DB: 2 Gaps: 1

09-833799-13B (1-57) x US-08-346-128-5 (1-2892)

QY 13 ProGlySerCysProIleIleLeuIleArgCysAlaMetLeuAsnProProAsnArgCys 32
|||
Db 316 CCAGGCACTTGCCCG-----TCGCTCCGTCCGACTCCCGGCTGC 278

QY 33 LeuLysAspThrAspCysProGlyIleLysLysCys 44
|||
Db 277 AGCAGCGACCTCGTTGTCCCGCCCGTGATCGTTGC 242

RESULT 63

US-08-532-384-1/c
; Sequence 1, Application US/08532384
; Patent No. 6083904
; GENERAL INFORMATION:
; APPLICANT: Artavanis-Tsakonas, S. et al.
; TITLE OF INVENTION: Therapeutic And Diagnostic Methods
; TITLE OF INVENTION: And Compositions Based On No. 6083904ch Proteins And
; TITLE OF INVENTION: Nucleic Acids
; NUMBER OF SEQUENCES: 21
; CORRESPONDENCE ADDRESSES:
; ADDRESSEE: Pennie & Edmonds
; STREET: 1155 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: U.S.A.
; ZIP: 10036
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/532,384
; FILING DATE:
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/083,590
; FILING DATE: 25-JUN-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Misrock, S. Leslie
; REGISTRATION NUMBER: 18,872
; REFERENCE/DOCKET NUMBER: 7326-015
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212 790-9090
; TELEFAX: 212 8698864/9741
; TELEX: 66141 PENNIE
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2892 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: unknown
; MOLECULE TYPE: cDNA
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 142..2640
US-08-532-384-1

Alignment Scores:
Pred. No.: 244 Length: 2892
Score: 61.50 Matches: 12
Percent Similarity: 43.75% Conservative: 2
Best Local Similarity: 37.50% Mismatches: 11
Query Match: 18.81% Indels: 7
DB: 3 Gaps: 1

09-833799-13B (1-57) x US-08-532-384-1 (1-2892)

QY 13 ProGlySerCysProIleIleLeuIleArgCysAlaMetLeuAsnProProAsnArgCys 32
|||
Db 316 CCAGGCACTTGCCCG-----TCGCTCCGTCCGACTCCCGGCTGC 278

QY 33 LeuLysAspThrAspCysProGlyIleLysLysCys 44
|||
Db 277 AGCAGCGACCTCGTTGTCCCGCCCGTGATCGTTGC 242

RESULT 64
US-08-893-828-5/c
; Sequence 5, Application US/08893828
; Patent No. 6090922
; GENERAL INFORMATION:
; APPLICANT: Artavanis-Tsakonas, Spyridon
; APPLICANT: Muskavitch, Marc A.T.

APPLICANT: Fehon, Richard G.
APPLICANT: Rebay, Ilaria
APPLICANT: Blaumueller, Cristine M.
APPLICANT: Shepard, Scott B.
TITLE OF INVENTION: HUMAN NOTCH AND DELTA, BINDING DOMAINS
TITLE OF INVENTION: IN TOPORYTHMIC PROTEINS, AND METHODS BASED THEREON
NUMBER OF SEQUENCES: 34
CORRESPONDENCE ADDRESS:
ADDRESSEE: PENNIE & EDMONDS
STREET: 1155 Avenue of the Americas
CITY: New York
STATE: NY
COUNTRY: USA
ZIP: 10036-2711
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/893,828
FILING DATE: 11-JUL-1997
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Mistrock, S. Leslie
REGISTRATION NUMBER: 18,872
REFERENCE/DOCKET NUMBER: 7326-050
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 790-9090
TELEFAX: (212) 869-8864/9741
INFORMATION FOR SEQ ID NO: 5:
SEQUENCE CHARACTERISTICS:
LENGTH: 2892 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: unknown
MOLECULE TYPE: cDNA
FEATURE:
NAME/KEY: CDS
LOCATION: 142..2640
US-08-893-828-5

Alignment Scores:
Pred. No.: 244 Length: 2892
Score: 61.50 Matches: 12
Percent Similarity: 43.75% Conservative: 2
Best Local Similarity: 37.50% Mismatches: 11
Query Match: 18.81% Indels: 7
DB: 3 Gaps: 1
09-833799-13B (1-57) x US-08-893-828-5 (1-2892)
QY 13 ProGlySerCysProIleIleLeuIleArgCysAlaMetLeuAsnProProAsnArgCys 32
Db 316 CCAGGCACTTGCCCG-----TCGCTCCGTCGCGACTCCCGCTGC 278
QY 33 LeuLysAspThrAspCysProGlyIleLysLysCys 44
Db 277 AGCAGCGACCTCGTTGTCGCCCGCCCGTGATCGTTGC 242
RESULT 65
US-08-747-887-4
; Sequence 4, Application US/08747887
; Patent No. 5853734
; GENERAL INFORMATION:
; APPLICANT: Chang, Yuan
; APPLICANT: Bohenzky, Roy A.
; APPLICANT: Russo, James J.
; APPLICANT: Edelman, Isidore S.
; APPLICANT: Moore, Patrick S.
; TITLE OF INVENTION: Glycoprotein L And Glycoprotein M From Kaposi's
; TITLE OF INVENTION: Sarcoma-Associated Herpesvirus, DNA Encoding
; TITLE OF INVENTION: Same And Uses Thereof

NUMBER OF SEQUENCES: 23
CORRESPONDENCE ADDRESS:
ADDRESSEE: Cooper & Dunham LLP
STREET: 1185 Avenue of the Americas
CITY: New York
STATE: New York
COUNTRY: U.S.A.
ZIP: 10036
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/747,887
FILING DATE:
CLASSIFICATION: 530
ATTORNEY/AGENT INFORMATION:
NAME: White, John P.
REGISTRATION NUMBER: 28,678
REFERENCE/DOCKET NUMBER: 52429
TELECOMMUNICATION INFORMATION:
TELEPHONE: 212-278-0400
TELEFAX: 212-391-0525
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 631 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
HYPOTHETICAL: N
ANTI-SENSE: N
US-08-747-887-4
Alignment Scores:
Pred. No.: 36.2 Length: 631
Score: 61.00 Matches: 18
Percent Similarity: 45.61% Conservative: 8
Best Local Similarity: 31.58% Mismatches: 22
Query Match: 18.65% Indels: 9
DB: 2 Gaps: 2
09-833799-13B (1-57) x US-08-747-887-4 (1-631)
QY 1 AlaGlnGluProValLysGlyProValSerThrLysProGlySerCysProIleIleLeu 20
Db 475 GCACGGGGTCCACAGGGCGGTCCGGTGGCATAACAGGCGG-----513
QY 21 IleArgCysAlaMetLeuAsnProPro----AsnArg-CysLeuLysAspThrAspCysPr 39
Db 514 GTGAGGGCCCCCTGGGTCTGTCCGCTCGAACAACAGGGTGTGTGAACAACAGGTTGCCA 573
QY 39 oGlyIleLysLysCysCysGluGlySerCysGlyMetAlaCysPheVal 55
Db 574 AGGCCGCGAATACCCCTCTGCACGCTGTGTGACGCTGGTGGTGTATGCTC 622
RESULT 66
US-08-728-323A-4
; Sequence 4, Application US/08728323A
; Patent No. 5948676
; GENERAL INFORMATION:
; APPLICANT: Chang, Yuan
; APPLICANT: Bohenzky, Roy A.
; APPLICANT: Russo, James J.
; APPLICANT: Edelman, Isidore S.
; APPLICANT: Moore, Patrick S.
; TITLE OF INVENTION: Immediate Early Protein From Kaposi's
; TITLE OF INVENTION: Sarcoma-Associated Herpesvirus, DNA
; TITLE OF INVENTION: Encoding Same And Uses Thereof
; NUMBER OF SEQUENCES: 21
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Cooper & Dunham LLP

/ STREET: 1185 Avenue of the Americas
/ CITY: New York
/ STATE: New York
/ COUNTRY: U.S.A.
/ ZIP: 10036
/ COMPUTER READABLE FORM:
/ MEDIUM TYPE: Floppy disk
/ COMPUTER: IBM PC compatible
/ OPERATING SYSTEM: PC-DOS/MS-DOS
/ SOFTWARE: PatentIn Release #1.0, Version #1.30
/ CURRENT APPLICATION DATA:
/ APPLICATION NUMBER: US/08/728,323A
/ FILING DATE:
/ CLASSIFICATION: 435
/ ATTORNEY/AGENT INFORMATION:
/ NAME: White, John P.
/ REGISTRATION NUMBER: 28,678
/ REFERENCE/DOCKET NUMBER: 0575/52268/JPW/MSK/SKS
/ TELECOMMUNICATION INFORMATION:
/ TELEPHONE: 212-278-0400
/ TELEFAX: 212-391-0525
/ INFORMATION FOR SEQ ID NO: 4:
/ SEQUENCE CHARACTERISTICS:
/ LENGTH: 631 base pairs
/ TYPE: nucleic acid
/ STRANDEDNESS: single
/ TOPOLOGY: linear
/ MOLECULE TYPE: DNA (genomic)
/ HYPOTHETICAL: N
/ ANTI-SENSE: N
/ US-08-728-323A-4

Alignment Scores:
Pred. No.: 36.2 Length: 631
Score: 61.00 Matches: 18
Percent Similarity: 45.61% Conservative: 8
Best Local Similarity: 31.58% Mismatches: 22
Query Match: 18.65% Indels: 9
DB: 2 Gaps: 2

09-833799-13B (1-57) x US-08-728-323A-4 (1-631)

QY 1 AlaGlnGluProValSerThryProValSerThryProGlySerCysProIleIleLeu 20
Db 475 GCACGGGGTCCCGGGCGGTCGGTGGCATAACAGGCCG----- 513
QY 21 IleArgCysAlaMetLeuAsnProPro---AsnArg-CysLeuLysAspThrAspCysPr 39
Db 514 GTGAGGGCCCCCTGGGTGTCTCCGCTGGAAACAGGGTGCTGTGAAACAAACAGGTTGCCA 573
QY 39 oGlyIleLysLysCysCysGluGlySerCysGlyMetAlaCysPheVal 55
Db 574 AGGCCCGGAATACCCCTCTGCACGCTGCTGTGGACGTGGGTGTATGCTC 622

RESULT 67
US-08-998-416-149/c
; Sequence 149, Application US/08998416
; Patent No. 6239264
; GENERAL INFORMATION:
; APPLICANT: Philippsen, Peter
; APPLICANT: Pohlmann, Rainer
; APPLICANT: Steiner, Sabine
; APPLICANT: Mohr, Christine
; APPLICANT: Wendland, Jurgen
; APPLICANT: Knechtle, Philipp
; APPLICANT: Rebischung, Corinne
; TITLE OF INVENTION: GENOMIC DNA SEQUENCES OF ASHBYA GOSSYPII
; TITLE OF INVENTION: AND USES THEREOF
; NUMBER OF SEQUENCES: 1152
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: No. 6239264artis Corporation
; STREET: 3054 Cornwallis Road
; CITY: Research Triangle Park

/ STATE: No. 6239264th Carolina
/ COUNTRY: USA
/ ZIP: 27709
/ COMPUTER READABLE FORM:
/ MEDIUM TYPE: Floppy disk
/ COMPUTER: IBM PC compatible
/ OPERATING SYSTEM: PC-DOS/MS-DOS
/ SOFTWARE: PatentIn Release #1.0, Version #1.30
/ CURRENT APPLICATION DATA:
/ APPLICATION NUMBER: US/08/998,416
/ FILING DATE: 24-DEC-1997
/ CLASSIFICATION: 435
/ PRIOR APPLICATION DATA:
/ APPLICATION NUMBER: CH 0016/97
/ FILING DATE: 31-DEC-1996
/ ATTORNEY/AGENT INFORMATION:
/ NAME: Meigs, J. Timothy
/ REGISTRATION NUMBER: 38,241
/ REFERENCE/DOCKET NUMBER: PF/5-30306/A/CGC1976
/ TELECOMMUNICATION INFORMATION:
/ TELEPHONE: 919-541-8587
/ TELEFAX: 919-541-8689
/ INFORMATION FOR SEQ ID NO: 149:
/ SEQUENCE CHARACTERISTICS:
/ LENGTH: 653 base pairs
/ TYPE: nucleic acid
/ STRANDEDNESS: single
/ TOPOLOGY: linear
/ MOLECULE TYPE: DNA (genomic)
/ ORIGINAL SOURCE:
/ ORGANISM: PAG1057RP
/ US-08-998-416-149

Alignment Scores:
Pred. No.: 37.9 Length: 653
Score: 61.00 Matches: 11
Percent Similarity: 63.64% Conservative: 3
Best Local Similarity: 50.00% Mismatches: 8
Query Match: 18.65% Indels: 0
DB: 4 Gaps: 0

09-833799-13B (1-57) x US-08-998-416-149 (1-653)

QY 7 GlyProValSerThryProGlySerCysProIleIleLeuIleArgCysAlaMetLeu 26
Db 275 GGACCCCTGTCCACACACCGTGCTCCTGCCCTCGTTGTTGCTCCGCTGCTGCTA 216
QY 27 AsnPro 28
Db 215 AACCCA 210

RESULT 68
US-09-056-556-191
; Sequence 191, Application US/09056556
; Patent No. 6350456
; GENERAL INFORMATION:
; APPLICANT: Reed, Steven G.
; APPLICANT: Skeiky, Yasir A.W.
; APPLICANT: Dillon, Davin C.
; TITLE OF INVENTION: COMPOUNDS AND METHODS FOR THE PREVENTION AND
; NUMBER OF SEQUENCES: 241
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SEED and BERRY LLP
; STREET: 6300 Columbia Center, 701 Fifth Avenue
; CITY: Seattle
; STATE: Washington
; COUNTRY: USA
; ZIP: 98104-7092
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30

TREA

;;
;; CURRENT APPLICATION DATA:
;; APPLICATION NUMBER: US/09/056,556
;; FILING DATE: 07-APR-1998
;; CLASSIFICATION:
;; ATTORNEY/AGENT INFORMATION:
;; NAME: Maki, David J.
;; REGISTRATION NUMBER: 31,392
;; REFERENCE/DOCKET NUMBER: 210121.457
;; TELECOMMUNICATION INFORMATION:
;; TELEPHONE: (206) 622-4900
;; TELEFAX: (206) 682-6031
;; INFORMATION FOR SEQ ID NO: 191:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 1055 base pairs
;; TYPE: nucleic acid
;; STRANDEDNESS: single
;; TOPOLOGY: linear
;; US-09-056-556-191

Alignment Scores: 72.4 Length: 1055
Pred. No.: 61.00 Matches: 20
Score: 32.10% Conservative: 6
Percent Similarity: 24.69% Mismatches: 19
Best Local Similarity: 18.65% Indels: 36
Query Match: 4
DB:

09-833799-13B (1-57) x US-09-056-556-191 (1-1055)

QY 3 GluProValLysGlyProValSer-----ThrLysProGlySer 15
Db 275 AAACCCGCGCCAGGCGCGCTGGACGGGCTCATGGCAGCGAAATTAGAAAACCCGGGATAT 334
QY 16 CysProIleIleLeuIleArgCysAlaMetLeuAsnProProAsnArgCysLeu----- 33
Db 335 TGTCGCGGATGTGCATACGATGCTGAGTGCTTGGT-----GGTTCGTGTTTAGCCATT 388
QY 34 -----LysAspThrAspCys----- 38
Db 389 GAGTGTGGATGTGTGAGACCCCTGGCCTGGAAGGGGACACCGTGCTTTTGCCCTCTTGTC 448
QY 39 -----ProGlyIleLysLysCysCysGluGlySer 48
Db 449 CGCCTTTGCCCGCCGACGCGGTGGTGGCGAAACGGGCTGAGTCGGAATGCTCGCGCGGT 508
QY 49 Cys 49
Db 509 TGT 511

RESULT 69
US-09-072-596-186
; Sequence 186, Application US/09072596
; Patent No. 6458366
; GENERAL INFORMATION:
; APPLICANT: Reed, Steven G.
; APPLICANT: Skeiky, Yasir A.W.
; APPLICANT: Dillon, Davin C.
; APPLICANT: Campos-Neto, Antonia
; APPLICANT: Houghton, Raymond
; APPLICANT: Vedvick, Thomas S.
; APPLICANT: Twardzik, Daniel R.
; APPLICANT: Lodes, Michael J.
; APPLICANT: Hendrickson, Ronald C.
; TITLE OF INVENTION: COMPOUNDS AND METHODS FOR DIAGNOSIS OF TUBERCULOSIS
; NUMBER OF SEQUENCES: 350
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SEED and BERRY LLP
; STREET: 6300 Columbia Center, 701 Fifth Avenue
; CITY: Seattle
; STATE: Washington
; COUNTRY: USA
; ZIP: 98104-7092
; COMPUTER READABLE FORM:

;;
;; MEDIUM TYPE: Floppy disk
;; COMPUTER: IBM PC compatible
;; OPERATING SYSTEM: PC-DOS/MS-DOS
;; SOFTWARE: PatentIn Release #1.0, Version #1.30
;; CURRENT APPLICATION DATA:
;; APPLICATION NUMBER: US/09/072,596
;; FILING DATE: 05-MAY-1998
;; CLASSIFICATION:
;; ATTORNEY/AGENT INFORMATION:
;; NAME: Maki, David J.
;; REGISTRATION NUMBER: 31,392
;; REFERENCE/DOCKET NUMBER: 210121.417C9
;; TELECOMMUNICATION INFORMATION:
;; TELEPHONE: (206) 622-4900
;; TELEFAX: (206) 682-6031
;; INFORMATION FOR SEQ ID NO: 186:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 1055 base pairs
;; TYPE: nucleic acid
;; STRANDEDNESS: single
;; TOPOLOGY: linear
;; US-09-072-596-186

Alignment Scores: 72.4 Length: 1055
Pred. No.: 61.00 Matches: 20
Score: 32.10% Conservative: 6
Percent Similarity: 24.69% Mismatches: 19
Best Local Similarity: 18.65% Indels: 36
Query Match: 4
DB:

09-833799-13B (1-57) x US-09-072-596-186 (1-1055)

QY 3 GluProValLysGlyProValSer-----ThrLysProGlySer 15
Db 275 AAACCCGCGCCAGGCGCGCTGGACGGGCTCATGGCAGCGAAATTAGAAAACCCGGGATAT 334
QY 16 CysProIleIleLeuIleArgCysAlaMetLeuAsnProProAsnArgCysLeu----- 33
Db 335 TGTCGCGGATGTGCATACGATGCTGAGTGCTTGGT-----GGTTCGTGTTTAGCCATT 388
QY 34 -----LysAspThrAspCys----- 38
Db 389 GAGTGTGGATGTGTGAGACCCCTGGCCTGGAAGGGGACACCGTGCTTTTGCCCTCTTGTC 448
QY 39 -----ProGlyIleLysLysCysCysGluGlySer 48
Db 449 CGCCTTTGCCCGCCGACGCGGTGGTGGCGAAACGGGCTGAGTCGGAATGCTCGCGCGGT 508
QY 49 Cys 49
Db 509 TGT 511

RESULT 70
US-09-615-192A-172/c
; Sequence 172, Application US/09615192A
; Patent No. 6410718
; GENERAL INFORMATION:
; APPLICANT: Bloksberg, Leonard N.
; APPLICANT: Havukkala, Ilkka
; TITLE OF INVENTION: Materials and Methods for the
; TITLE OF INVENTION: Modification of Plant Lignin Content
; FILE REFERENCE: 11000.1003C4U
; CURRENT APPLICATION NUMBER: US/09/615,192A
; CURRENT FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 08/975,316
; PRIOR FILING DATE: 1997-11-21
; PRIOR APPLICATION NUMBER: US 08/713,000
; PRIOR FILING DATE: 1996-09-11
; PRIOR APPLICATION NUMBER: US 09/169,789
; PRIOR FILING DATE: 1998-10-09
; NUMBER OF SEQ ID NOS: 405
; SOFTWARE: FastSEQ for Windows Version 3.0

```

US-08-858-207A-161/c
; Sequence 161, Application US/08858207A
; Patent No. 6348328
; GENERAL INFORMATION:
; APPLICANT: Black, Michael
; APPLICANT: Hodgson, John
; APPLICANT: Knowles, David
; APPLICANT: Nicholas, Richard
; APPLICANT: Stodola, Robert
; TITLE OF INVENTION: No. 6348328e1 Compounds
; NUMBER OF SEQUENCES: 552
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SmithKline Beecham Corporation
; STREET: 709 Swedeland Road
; CITY: King of Prussia
; STATE: PA
; COUNTRY: USA
; ZIP: 19406-0939
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSEQ for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/858,207A
; FILING DATE: 09-MAY-1997
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 60/017670
; FILING DATE: 14-MAY-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Gimmi, Edward R
; REGISTRATION NUMBER: 38,891
; REFERENCE/DOCKET NUMBER: P50475
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 610-270-4478
; TELEFAX: 610-270-5090
; TELEX:
; INFORMATION FOR SEQ ID NO: 161:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1089 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
US-08-858-207A-161

Alignment Scores:
Pred. No.: 102 Length: 1089
Score: 60.00 Matches: 13
Percent Similarity: 43.75% Conservative: 8
Best Local Similarity: 27.08% Mismatches: 11
Query Match: 18.35% Indels: 16
DB: 4 Gaps: 1

09-833799-13B (1-57) x US-08-858-207A-161 (1-1089)
QY 8 ProValSerThrLysProGlySerCysProIleIleLeuIleArgCysAlaMetLeuAsn 27
Db 912 CCAATAGCTTCTGCCAGATGGTCTTGCCCAATAATTCATGAAAAATCGGAGTGAGCATG 853
QY 28 ProProAsnArgCysLeu-----
Db 852 GCGCCTGAAAAAGCACTCAATGTGTGACACAAGAGGTTAAGCAGACAAATAAAGGTACT 793
QY 34 -----LysAspThrAspCysPro 39
Db 792 AGCAACAAGGAAAAAGACTGCGCCT 769

RESULT 73
US-09-282-147-37/c
; Sequence 37, Application US/09282147
; Patent No. 6274147
; GENERAL INFORMATION:

```

APPLICANT: VAKHARIA, Vikram
APPLICANT: YAO, Kun
TITLE OF INVENTION: METHOD FOR GENERATING NONPATHOGENIC, INFECTIOUS
TITLE OF INVENTION: PANCREATIC NECROSIS VIRUS (IPNV) FROM SYNTHETIC RNA
TITLE OF INVENTION: TRANSCRIPTS
FILE REFERENCE: 8288-9023
CURRENT APPLICATION NUMBER: US/09/282,147
CURRENT FILING DATE: 1999-03-31
EARLIER APPLICATION NUMBER: US/60/080,278
EARLIER FILING DATE: 1998-03-31
EARLIER APPLICATION NUMBER: PCT/US97/12955
EARLIER FILING DATE: 1998-03-31
NUMBER OF SEQ ID NOS: 51
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 37
LENGTH: 2783
TYPE: DNA
ORGANISM: Infectious pancreatic necrosis virus
US-09-282-147-37

Alignment Scores:
Pred. No.: 360 Length: 2783
Score: 60.00 Matches: 15
Percent Similarity: 41.67% Conservative: 5
Best Local Similarity: 31.25% Mismatches: 26
Query Match: 18.35% Indels: 2
DB: 4 Gaps: 1

09-833799-13B (1-57) x US-09-282-147-37 (1-2783)

QY 8 ProValSerThrLysProGlySerCysProIleLeuLeuLeuArgCysAlaMetLeuAsn 27
Db 1376 CCTGTCAGGAGGTAGTACATCATGGCCTGCATGTGTGGAGTGACGTTGGCCTCGC 1317
QY 28 ProProAsnArgCysLeuLysAspThrAspCysProGlyIleLysLysCysCys----- 45
Db 1316 CCTTTCTAGGCTATTGAGTACACAGTGTGTCTCTGCAGTATGATGTTGTCTGCAT 1257
QY 46 GluGlySerCysGlyMetAlaCys 53
Db 1256 AGATCATCTAGGTCCTCGTTGT 1233

RESULT 74
US-09-543-084A-29/c
Sequence 29, Application US/09543084A
Patent No. 6361988
GENERAL INFORMATION:
APPLICANT: Frances H. Arnold
APPLICANT: Zhixin Shao
APPLICANT: Huimin Zhao
APPLICANT: Lorraine J. Giver
TITLE OF INVENTION: Recombination of Polynucleotide
TITLE OF INVENTION: Sequences Using Defined or Random Primer Sequences
NUMBER OF SEQUENCES: 36
CORRESPONDENCE ADDRESS:
ADDRESSEE: Oppenheimer Wolff & Donnelly LLP
STREET: 2029 Century Park East, Suite 3800
CITY: Los Angeles
STATE: CA
COUNTRY: USA
ZIP: 90067
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: Windows 98
SOFTWARE: Word 97
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/543,084A
FILING DATE: April 4, 2000
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 09/381,935
APPLICATION NUMBER: 60/041,666

FILING DATE: March 25, 1997
APPLICATION NUMBER: 60/045,211
FILING DATE: April 30, 1997
APPLICATION NUMBER: 60/046,256
FILING DATE: May 12, 1997
APPLICATION NUMBER: 08/905,359
FILING DATE: August 4, 1997
ATTORNEY/AGENT INFORMATION:
NAME: Oldenkamp, David J.
REGISTRATION NUMBER: 29,421
REFERENCE/DOCKET NUMBER: 330187-89
TELEPHONE: (310) 788-5000
TELEFAX: (310) 788-5100
INFORMATION FOR SEQ ID NO: 29:
SEQUENCE CHARACTERISTICS:
LENGTH: 3222 nucleotides
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: polynucleotide
US-09-543-084A-29

Alignment Scores:
Pred. No.: 439 Length: 3222
Score: 60.00 Matches: 15
Percent Similarity: 40.00% Conservative: 5
Best Local Similarity: 30.00% Mismatches: 16
Query Match: 18.35% Indels: 14
DB: 4 Gaps: 3

09-833799-13B (1-57) x US-09-543-084A-29 (1-3222)

QY 16 CysProIleLeuLeuLeuArgCysAlaMetLeuAsnProProAsnArgCysLeuLysAsp 35
Db 1209 TGCCCGATC-----TCGATGATCGGTGCGCGACAGCGCGCGCTCG 1165
QY 36 ThrAsp---CysProGlyIleLysLysCys-----CysGlu 46
Db 1164 ACGTCGTAGCGCGCGCACCTTGAGTGCATCCGGTAGAGCGTGGCGCGCTGCCAC 1105
QY 47 GlySerCysGlyMetAlaCysPheValPro 56
Db 1104 GGAAGTGCAGGTGGCCAGCACCATCCG 1075

RESULT 75
US-09-543-084A-26/c
Sequence 26, Application US/09543084A
Patent No. 6361988
GENERAL INFORMATION:
APPLICANT: Frances H. Arnold
APPLICANT: Zhixin Shao
APPLICANT: Huimin Zhao
APPLICANT: Lorraine J. Giver
TITLE OF INVENTION: Recombination of Polynucleotide
TITLE OF INVENTION: Sequences Using Defined or Random Primer Sequences
NUMBER OF SEQUENCES: 36
CORRESPONDENCE ADDRESS:
ADDRESSEE: Oppenheimer Wolff & Donnelly LLP
STREET: 2029 Century Park East, Suite 3800
CITY: Los Angeles
STATE: CA
COUNTRY: USA
ZIP: 90067
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: Windows 98
SOFTWARE: Word 97
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/543,084A
FILING DATE: April 4, 2000
CLASSIFICATION: 435

PRIOR APPLICATION DATA:
APPLICATION NUMBER: 09/381,935
FILING DATE: March 25, 1997
APPLICATION NUMBER: 60/041,666
FILING DATE: April 30, 1997
APPLICATION NUMBER: 60/045,211
FILING DATE: May 12, 1997
APPLICATION NUMBER: 60/046,256
FILING DATE: August 4, 1997
APPLICATION NUMBER: 08/905,359
FILING DATE: August 4, 1997
ATTORNEY/AGENT INFORMATION:
NAME: Oldenkamp, David J.
REGISTRATION NUMBER: 29,421
REFERENCE/DOCKET NUMBER: 330187-89
TELEPHONE: (310) 788-5000
TELEFAX: (310) 788-5100
INFORMATION FOR SEQ ID NO: 26:
SEQUENCE CHARACTERISTICS:
LENGTH: 4050 nucleotides
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: polynucleotide
US-09-543-084A-26

Alignment Scores:
Pred. No.: 598 Length: 4050
Score: 60.00 Matches: 15
Percent Similarity: 40.00% Conservative: 5
Best Local Similarity: 30.00% Mismatches: 16
Query Match: 18.35% Indels: 14
DB: 4 Gaps: 3

09-833799-13B (1-57) x US-09-543-084A-26 (1-4050)
QY 16 CysProIleIleLeuIleArgCysAlaMetLeuAsnProProAsnArgCysLeuLysAsp 35
Db 2083 TGCCCGATC-----TCGATGATCGGGTCGCCGACCGAGCGCGCCCTCG 2039
QY 36 ThrAsp---CysProGlyIleLysLysCys-----CysGlu 46
Db 2038 ACGTCGTAGCGCGCGGCACCTTGAGTGATCGCGGTAGAGCGTGGCGCGCCCTGCCAC 1979
QY 47 GlySerCysGlyMetAlaCysPheValPro 56
Db 1978 GGGAAGTGGGGTTGGCCAGCACCATCCCG 1949

Search completed: February 15, 2003, 21:14:41
Job time : 101 secs

GenCore version 5.1.4_p5_4578
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 15, 2003, 21:53:32 ; Search time 1184 Seconds
(without alignments)
2817.796 Million cell updates/sec

Title: 09-833799-13A
Perfect score: 206
Sequence: 1 AATTCGAGCTCGGTACCATA.....GTTTCGTTCCACAATAATAG 206

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 16154066 seqs, 8097743376 residues
Total number of hits satisfying chosen parameters: 32308132

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 100 summaries

Database :

EST: *
1: em_estba: *
2: em_esthum: *
3: em_estin: *
4: em_estmu: *
5: em_estov: *
6: em_estpl: *
7: em_estro: *
8: em_htc: *
9: gb_est1: *
10: gb_est2: *
11: gb_htc: *
12: gb_est3: *
13: gb_est4: *
14: gb_est5: *
15: em_estfun: *
16: em_estom: *
17: gb_gss: *
18: em_gss_hum: *
19: em_gss_inv: *
20: em_gss_pln: *
21: em_gss_vrt: *
22: em_gss_fun: *
23: em_gss_mam: *
24: em_gss_mus: *
25: em_gss_other: *
26: em_gss_pro: *
27: em_gss_rod: *

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	126.2	61.3	388	12	BE772159 CM4-FT010
2	126.2	61.3	423	9	AI830872 wj61c08.x
3	126.2	61.3	438	10	AW001880 wt81d06.x
4	124.6	60.5	219	10	BE182652 RC3-HT064
5	124.6	60.5	340	13	BG987761 CM3-HT119
6	124.6	60.5	342	10	BE182993 CM4-HT065

C	80	120.8	58.6	344	10	AW844969
	81	120.2	58.3	391	12	BE772162
	82	120	58.3	298	12	BF837690
	83	120	58.3	313	12	BF736294
	84	120	58.3	314	12	BF837677
C	85	119.8	58.2	431	12	BF837730
	86	116.6	56.6	470	12	BF913875
C	87	115.4	56.0	448	10	AW082007
C	88	114.6	55.6	289	12	BF834601
C	89	113.6	55.1	275	10	BE182650
C	90	113.6	55.1	275	12	BE711090
	91	113.6	55.1	393	10	BE183902
	92	113.6	55.1	393	12	BE772164
C	93	112.6	54.7	400	9	AI831796
C	94	112.6	54.7	411	9	AA586718
C	95	112.6	54.7	470	12	BF913869
C	96	112.6	54.7	478	9	AA527557
C	97	112.4	54.6	555	9	AA583567
	98	112.2	54.5	468	12	BF090376
	99	111.8	54.3	336	12	BF844574
C	100	110	53.4	190	10	BE181433

ALIGNMENTS

```
RESULT 1
BE772159
LOCUS BE772159 388 bp mRNA linear EST 20-SEP-2000
DEFINITION CM4-FT0104-230600-215-c06 FT0104 Homo sapiens cDNA, mRNA sequence.
ACCESSION BE772159
VERSION BE772159.1 GI:10225817
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 388)
Dias Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Briones,M.R.,
Nagai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.F.,
Goldman,G.H., Carvalho,A.F., Matsukuma,A., Baia,G.S., Simpson,D.H.,
Brunstein,A., deOliveira,P.S., Bucher,P., Jongeneel,C.V., O'Hare
,M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and
Simpson,A.J.
Shotgun sequencing of the human transcriptome with ORF expressed
sequence tags
Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
20202663
Contact: Simpson A.J.G.
Laboratory of Cancer Genetics
Ludwig Institute for Cancer Research
Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
Brazil
Tel: +55-11-2704922
Fax: +55-11-2707001
Email: asimpson@ludwig.org.br
This sequence was derived from the FAPESP/LICR Human Cancer Genome
Project. This entry can be seen in the following URL
(http://www.ludwig.org.br/scripts/gethtml2.pl?tl=&t2=CM4-FT0104-230
600-215-c06&t3=2000-06-23&t4=1)
Seq primer: puc 18 forward
High quality sequence start: 26
High quality sequence stop: 388.
Location/Qualifiers
1 .388
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_lib="FT0104"
/dev_stage="Adult"
```

```
FEATURES
source
1 .388
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_lib="FT0104"
/dev_stage="Adult"
/notes="Organ: prostate tumor; Vector: puc18; Site_1: SmaI;
Site_2: SmaI; A mini-library was made by cloning products
derived from ORESTES PCR (U.S. Letters Patent application
No. 196,716 - Ludwig Institute for Cancer Research)
```

```
profiles into the pUC 18 vector. Reverse transcription of
tissue mRNA and cDNA amplification were performed under
low stringency conditions."
BASE COUNT 88 a 115 c 93 g 92 t
ORIGIN
Query Match 61.3%; Score 126.2; DB 12; Length 388;
Best Local Similarity 81.6%; Pred. No. 8.8e-30;
Matches 146; Conservative 0; Mismatches 33; Indels 0; Gaps 0;
QY 28 ATGCTCAAGAACCGTTAAAGGTCCTGTCTACTAAGCCAGGTTCTTGTCCTATTATCT 87
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 113 AAGCGCAAGAGCCAGTCAAGGTCCTCCAGTCTCCACTAAGCCTGGCTCCTCCCTATTATCT 172
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 88 TGATTCGTTGGCTATGTTAAACCCACCTAACCGTTGTTTGAAGGACACTGATTGTCCAG 147
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 173 TGATCCGGTGGCCCATGTTGAATCCCCCTAACCGCTGTTGAAAGATACTGACTGCCAG 232
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 148 GTATCAAAAAGTGTGTAAGGTTCTCGGGTATGGCTTGTTCGTTCCACAATAATAG 206
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 233 GAATCAAGAAAGTGTGTGAAGGCTCTTGGGGATGGCCTGTTTCGTTCCCGAGTGAGAG 291
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
```

```
RESULT 2
AI830872/c
LOCUS AI830872 423 bp mRNA linear EST 21-DEC-1999
DEFINITION w161c08.x1 NCI CGAP Lul9 Homo sapiens cDNA clone IMAGE:2407310 3,
similar to gb:S58717_rnal ELAFIN PRECURSOR (HUMAN); contains element
MER22 repetitive element ;, mRNA sequence.
ACCESSION AI830872
VERSION AI830872.1 GI:5451543
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 423)
NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
Unpublished (1997)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: Christopher Moskaluk, M.D., Ph.D., Michael R.
Emmert-Buck, M.D., Ph.D.
cDNA Library Preparation: M. Bento Soares, Ph.D.
cDNA Library Arrayed by: Greg Lennon, Ph.D.
DNA Sequencing by: Washington University Genome Sequencing Center
Clone distribution: NCI-CGAP clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
www-bio.llnl.gov/bbrp/image/image.html
Insert Length: 534 Std Error: 0.00
Seq primer: -40UP from Gibco.
Location/Qualifiers
1 .423
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:2407310"
/clone_lib="NCI_CGAP_Lul9"
/tissue type="squamous cell carcinoma, poorly
differentiated (4 pooled tumors, including primary and
metastatic)"
/dev_stage="adult"
/lab_host="DH10B (phage-resistant)"
/notes="Organ: lung; Vector: pT7T3D-Pac (Pharmacia) with a
modified polylinker; 1st strand cDNA was prepared from
pooled lung tumor tissue, and was then primed with a Not I
- oligo(dT) primer. Double-stranded cDNA was ligated to
Eco RI adaptors (Pharmacia), digested with Not I and
cloned into the Not I and Eco RI sites of the modified
pT7T3 vector. Library went through one round of
normalization. Library constructed by Bento Soares and M.
Fatima Bonaudo."
```

```
FEATURES
source
1 .423
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:2407310"
/clone_lib="NCI_CGAP_Lul9"
/tissue type="squamous cell carcinoma, poorly
differentiated (4 pooled tumors, including primary and
metastatic)"
/dev_stage="adult"
/lab_host="DH10B (phage-resistant)"
/notes="Organ: lung; Vector: pT7T3D-Pac (Pharmacia) with a
modified polylinker; 1st strand cDNA was prepared from
pooled lung tumor tissue, and was then primed with a Not I
- oligo(dT) primer. Double-stranded cDNA was ligated to
Eco RI adaptors (Pharmacia), digested with Not I and
cloned into the Not I and Eco RI sites of the modified
pT7T3 vector. Library went through one round of
normalization. Library constructed by Bento Soares and M.
Fatima Bonaudo."
```

```
BASE COUNT      106 a      91 c      125 g      101 t
ORIGIN
Query Match      61.3%; Score 126.2; DB 9; Length 423;
Best Local Similarity 81.6%; Pred. No. 9.1e-30;
Matches 146; Conservative 0; Mismatches 33; Indels 0; Gaps 0;

Qy 28 ATGCTCAAGAACCCAGTTAAAGTCTGTGTCTACTAAGCCAGGTTCTTGTCTCATTATCT 87
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 361 AAGCGCAAGAGCCAGTCAAGGTCAGTCTCCAATAAGCCTGGCTCCTGCCCATATCT 302

Qy 88 TGATTCGTTGCGCTATGTTAAACCCACCTAACCGTTGTTGAAGGACACTGATGTCCAG 147
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 301 TGATCCGGTGGCCCATGTTGAATCCCCCTAACCGGTGTTGAAGAACTGACTGCCAG 242

Qy 148 GTATCAAAAAGTCTGTGAAGGTTCTCCTGCGGTATGGCTGTTTCGTTCCACAATAATAG 206
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 241 GAATCAAGAAGTCTGTGAAGGTTCTTGCGGGATGGCCTGTTTCGTTCCCAAGTGAGAG 183

RESULT 3
AW001880/c
LOCUS
DEFINITION
IMAGE:2513867 3' Similar to gb:S58717 rnal ELAFIN PRECURSOR (HUMAN)
);contains element MER28 repetitive element ;, mRNA sequence.
ACCESSION
AW001880
VERSION
AW001880.1 GI:5848796
KEYWORDS
EST.
SOURCE
human.
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 438)
REFERENCE
1 (bases 1 to 438)
AUTHORS
NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
TITLE
National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
JOURNAL
Unpublished (1997)
COMMENT
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
This clone is available royalty-free through LLNL; contact the
IMAGE Consortium (info@image.llnl.gov) for further information.
Insert Length: 552 Std Error: 0.00
Seq primer: -40UP from Gibco.
FEATURES
source
1..438
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:2513867"
/clone_lib="Soares_thymus_NHFT"
/dev_stage="fetal"
/lab_host="DH10B (phage-resistant)"
/note="Organ: thymus, pooled; Vector: pT7T3D-Pac
(Pharmacia) with a modified polylinker; Site 1: Not I;
Site 2: Eco RI; 1st strand cDNA was primed with a Not I -
oligo(dT) primer [5',
TGTTACCAATCTGAAGTGGAGCGCGCAACGTTTTTTTTTTTTTTT 3'],
double-stranded cDNA was ligated to Eco RI adaptors
(Pharmacia), digested with Not I and cloned into the Not I
and Eco RI sites of the modified pT7T3 vector. Library
went through one round of normalization. Library
constructed by Bento Soares and M. Fatima Bonaldo."

BASE COUNT      106 a      92 c      125 g      115 t
ORIGIN
Query Match      61.3%; Score 126.2; DB 10; Length 438;
Best Local Similarity 81.6%; Pred. No. 9.2e-30;
Matches 146; Conservative 0; Mismatches 33; Indels 0; Gaps 0;

Qy 28 ATGCTCAAGAACCCAGTTAAAGTCTGTGTCTACTAAGCCAGGTTCTTGTCTCATTATCT 87
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 377 AAGCGCAAGAGCCAGTCAAGGTCAGTTTCCACTAAGCCTGGCTCCTGCCCATATTT 318
```

```
Qy 88 TGATTCGTTGCGCTATGTTAAACCCACCTAACCGTTGTTGAAGGACACTGATGTCCAG 147
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 317 TGATCCGGTGGCCCATGTTGAATCCCCCTAACCGTCTGTTGAAAGATACTGACTGCCAG 258

Qy 148 GTATCAAAAAGTCTGTGAAGGTTCTCCTGCGGTATGGCTGTTTCGTTCCACAATAATAG 206
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 257 GAATCAAAAAGTCTGTGAAGGTTCTTTCGCGGATGGCCTGTTTCGTTCCCAAGTGAGAG 199

RESULT 4
BE182652
LOCUS
DEFINITION
RC3-HT0649-090500-011-h03 HT0649 Homo sapiens cDNA, mRNA sequence.
ACCESSION
BE182652
VERSION
BE182652.1 GI:8661828
KEYWORDS
EST.
SOURCE
human.
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 219)
REFERENCE
1 (bases 1 to 219)
AUTHORS
Dias Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Briones,M.R.,
Nagai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.F.,
Goldman,G.H., Carvalho,A.F., Matsukuma,A., Baia,G.S., Simpson,D.H.,
Brunstein,A., deOliveira,P.S., Bucher,P., Jongeneel,C.V., O'Hare
,M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and
Simpson,A.J.
TITLE
Shotgun sequencing of the human transcriptome with ORF expressed
sequence tags
JOURNAL
Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
MEDLINE
20202663
COMMENT
Contact: Simpson A.J.G.
Laboratory of Cancer Genetics
Ludwig Institute for Cancer Research
Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
Brazil
Tel: +55-11-2704922
Fax: +55-11-2707001
Email: asimpson@ludwig.org.br
This sequence was derived from the FAPESP/LICR Human Cancer Genome
Project. This entry can be seen in the following URL
(http://www.ludwig.org.br/scripts/gethtml2.pl?t1=&t2=RC3-HT0649-090
500-011-h03&t3=2000-05-09&t4=1)
Seq primer: puc 18 forward
High quality sequence start: 22
High quality sequence stop: 219.
FEATURES
Location/Qualifiers
1..219
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_lib="HT0649"
/dev_stage="Adult"
/note="Organ: head neck; Vector: puc18; Site 1: SmaI;
Site 2: SmaI; A mini-library was made by cloning products
derived from ORESTES PCR (U.S. Letters Patent application
No. 196,716 - Ludwig Institute for Cancer Research)
profiles into the pUC 18 vector. Reverse transcription of
tissue mRNA and cDNA amplification were performed under
low stringency conditions."
BASE COUNT      51 a      60 c      59 g      49 t
ORIGIN
Query Match      60.5%; Score 124.6; DB 10; Length 219;
Best Local Similarity 81.0%; Pred. No. 2.4e-29;
Matches 145; Conservative 0; Mismatches 34; Indels 0; Gaps 0;

Qy 28 ATGCTCAAGAACCCAGTTAAAGTCTGTGTCTACTAAGCCAGGTTCTTGTCTCATTATCT 87
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 22 AAGCGCAAGAGCCAGTCAAGGTCAGTCTCCACTAAGCCTGGCTCCTGCCCATATCT 81

Qy 88 TGATTCGTTGCGCTATGTTAAACCCACCTAACCGTTGTTGAAGGACACTGATGTCCAG 147
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 82 TGATCCGGTGGCCCATGTTGAATCCCCCTAACCGTCTGTTGAAAGATACTGACTGCCAG 141
```

QY 148 GTATCAAAAAGTCTGTGAAGGTTCTTGGGTATGGCTTGTTCGTTCCACAAATAATAG 206
|||||
Db 142 GAATCAAGAAGTCTGTGAAGGCTCTTGGGGATGGCCTGTTCTGTTCCCGAGTGAGAG 200
|||||

RESULT 5
BG987761/c
LOCUS CM3-HT1190-110101-624-h04 HT1190 Homo sapiens cDNA, mRNA sequence. EST 13-JUN-2001
DEFINITION
ACCESSION BG987761
VERSION BG987761.1 GI:14391831
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 340)
AUTHORS Dias Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Briones,M.R., Nagai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.F., Goldman,G.H., Carvalho,A.F., Matsukuma,A., Baia,G.S., Simpson,D.H., Brunstein,A., deOliveira,P.S., Bucher,P., Jongeneel,C.V., O'Hare ,M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and Simpson,A.J.
TITLE Shotgun sequencing of the human transcriptome with ORF expressed sequence tags
JOURNAL Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
MEDLINE 20202663
COMMENT Contact: Simpson A.J.G.
Laboratory of Cancer Genetics
Ludwig Institute for Cancer Research
Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP, Brazil
Tel: +55-11-2704922
Fax: +55-11-2707001
Email: asimpson@ludwig.org.br
This sequence was derived from the FAPESP/LICR Human Cancer Genome Project. This entry can be seen in the following URL (http://www.ludwig.org.br/scripts/gethtml2.pl?tl=CM3&t2=CM3-HT1190-110101-624-h04&t3=2001-01-11&t4=1)
Seq primer: puc 18 forward
High quality sequence start: 6
High quality sequence stop: 340.

FEATURES
source
1. .340
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_lib="HT1190"
/dev_stage="Adult"
/note="Organ: head neck; Vector: puc18; Site_1: SmaI;
Site_2: SmaI; A mini-library was made by cloning products derived from ORESTES PCR (U.S. Letters Patent application No. 196,716 - Ludwig Institute for Cancer Research) profiles into the pUC 18 vector. Reverse transcription of tissue mRNA and cDNA amplification were performed under low stringency conditions."
BASE COUNT 79 a 82 c 97 g 82 t
ORIGIN
Query Match 60.5%; Score 124.6; DB 13; Length 340;
Best Local Similarity 81.0%; Pred. No. 2.8e-29;
Matches 145; Conservative 0; Mismatches 34; Indels 0; Gaps 0;

QY 28 ATGCTCAAGAACCCAGTTAAAGGTCCTGTGTCTACTAAGCCAGGTTCTTGTCTTATATCT 87
|||||
Db 240 AAGCGCAAGAGCCAGTCAAGGTCAGTCTCCACTAAGCTGGCTCCTGCCCATATATCT 181
|||||

QY 88 TGATTCGTTGCGCTATGTTAAACCCACCTAACCGTTGTTGAAGGACACTGATGTCCAG 147
|||||
Db 180 TGATCCGTTGCGCCATGTTGAATCCCCCTAACCGCTGTTGAAGATAGTACTGACTGCCAG 121
|||||

QY 148 GTATCAAAAAGTCTGTGAAGGTTCTTGGGTATGGCTTGTTCGTTCCACAAATAATAG 206
|||||

Db 120 GAATCAAGAAGTCTGTGAAGGCTCTTGGGGATGGCCTGTTCTGTTCCCGAGTGAGAG 62
|||||

RESULT 6
BE182993/c
LOCUS CM4-HT0653-180400-146-a11 HT0653 Homo sapiens cDNA, mRNA sequence. EST 22-JUN-2000
DEFINITION
ACCESSION BE182993
VERSION BE182993.1 GI:8662169
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 342)
AUTHORS Dias Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Briones,M.R., Nagai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.F., Goldman,G.H., Carvalho,A.F., Matsukuma,A., Baia,G.S., Simpson,D.H., Brunstein,A., deOliveira,P.S., Bucher,P., Jongeneel,C.V., O'Hare ,M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and Simpson,A.J.
TITLE Shotgun sequencing of the human transcriptome with ORF expressed sequence tags
JOURNAL Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
MEDLINE 20202663
COMMENT Contact: Simpson A.J.G.
Laboratory of Cancer Genetics
Ludwig Institute for Cancer Research
Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP, Brazil
Tel: +55-11-2704922
Fax: +55-11-2707001
Email: asimpson@ludwig.org.br
This sequence was derived from the FAPESP/LICR Human Cancer Genome Project. This entry can be seen in the following URL (http://www.ludwig.org.br/scripts/gethtml2.pl?tl=CM4-HT0653-180400-146-a11&t3=2000-04-18&t4=1)
Seq primer: puc 18 forward
High quality sequence start: 21
High quality sequence stop: 342.

FEATURES
source
1. .342
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_lib="HT0653"
/dev_stage="Adult"
/note="Organ: head neck; Vector: puc18; Site_1: SmaI;
Site_2: SmaI; A mini-library was made by cloning products derived from ORESTES PCR (U.S. Letters Patent application No. 196,716 - Ludwig Institute for Cancer Research) profiles into the pUC 18 vector. Reverse transcription of tissue mRNA and cDNA amplification were performed under low stringency conditions."
BASE COUNT 80 a 82 c 100 g 79 t 1 others
ORIGIN
Query Match 60.5%; Score 124.6; DB 10; Length 342;
Best Local Similarity 81.0%; Pred. No. 2.8e-29;
Matches 145; Conservative 0; Mismatches 34; Indels 0; Gaps 0;

QY 28 ATGCTCAAGAACCCAGTTAAAGGTCCTGTGTCTACTAAGCCAGGTTCTTGTCTTATATCT 87
|||||
Db 261 AAGCGCAAGAGCCAGTCAAGGTCAGTCTCCACTAAGCCTGGCTCCTGCCCATATATCT 202
|||||

QY 88 TGATTCGTTGCGCTATGTTAAACCCACCTAACCGTTGTTGAAGGACACTGATGTCCAG 147
|||||
Db 201 TGATTCGTTGCGCCATGTTGAATCCCCCTAACCGCTGCTTGAAGATAGTACTGACTGCCAG 142
|||||

QY 148 GTATCAAAAAGTCTGTGAAGGTTCTTGGGTATGGCTTGTTCGTTCCACAAATAATAG 206
|||||

Db 141 GAATCAAGAAGTCTGTGAAGGCTCTTGGGGATGGCCTGTTCTGTTCCCGAGTGAGAG 83
|||||

RESULT 7
BE181546/c
LOCUS
DEFINITION CM0-HT0634-270300-308-f02 HT0634 Homo sapiens cDNA, mRNA sequence. EST 22-JUN-2000
ACCESSION BE181546
VERSION BE181546
KEYWORDS BE181546.1 GI:8660722
SOURCE EST.
ORGANISM human.
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 344)
Dias Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Briones,M.R.,
Nagai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.F.,
Goldman,G.H., Carvalho,A.F., Matsukuma,A., Baia,G.S., Simpson,D.H.,
Brunstein,A., deOliveira,P.S., Bucher,P., Jongeneel,C.V., O'Hare
,M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and
Simpson,A.J.
Shotgun sequencing of the human transcriptome with ORF expressed
sequence tags
Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
20202663
Contact: Simpson A.J.G.
Laboratory of Cancer Genetics
Ludwig Institute for Cancer Research
Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
Brazil
Tel: +55-11-2704922
Fax: +55-11-2707001
Email: asimpson@ludwig.org.br
This sequence was derived from the FAPESP/LICR Human Cancer Genome
Project. This entry can be seen in the following URL
(http://www.ludwig.org.br/scripts/gethtml2.pl?tl=&t2=CM0-HT0634-270
300-308-f02&t3=2000-03-27&t4=1)
Seq primer: puc 18 forward
High quality sequence start: 6
High quality sequence stop: 344.
FEATURES
Location/Qualifiers
1..344
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_lib="HT0634"
/dev_stage="Adult"
/note="Organ: head neck; Vector: puc18; Site_1: SmaI;
Site_2: SmaI; A mini-library was made by cloning products
derived from ORESTES PCR (U.S. Letters Patent application
No. 196,716 - Ludwig Institute for Cancer Research)
profiles into the pUC 18 vector. Reverse transcription of
tissue mRNA and cDNA amplification were performed under
low stringency conditions."
BASE COUNT 79 a 80 c 113 g 72 t
ORIGIN
Query Match 60.5%; Score 124.6; DB 10; Length 344;
Best Local Similarity 81.0%; Pred. No. 2.8e-29;
Matches 145; Conservative 0; Mismatches 34; Indels 0; Gaps 0;
QY 28 ATGCTCAAGAACCCAGTTAAAGGTCCTGTGCTACTAAGCCAGGTTCTTGCTCTATTATCT 87
Db 296 AAGCGCAAGAGCCAGTCAAAGGTCCTGCTCCACTAAGCCTGGCTCTGCCCCATTATCT 237
QY 88 TGATTCGTTGGCGTATGTTAAACCCACCTAACCGTTGTTTGAAGGACACTGATTGTCCAG 147
Db 236 TGATCCGGTGGCGCATGTTGAATCCCCCTAACCGTCTGCTTGAAGATACTGACTGCCAG 177
QY 148 GTATCAAAAAGTGCTGTGAAGGTTCCTCGGGTATGGCTTGTTCGTTCCACAATAATAG 206
Db 176 GAATCAAGAAGTGCTGTGAAGGCTCTTCGGGGATGGCCTGTTTCGTTCCCCAGTGAGAG 118
RESULT 8
BE182864
LOCUS
DEFINITION CM4-HT0652-150400-143-a07 HT0652 Homo sapiens cDNA, mRNA sequence.
ACCESSION BE182864
VERSION BE182864.1 GI:8662040
KEYWORDS EST.
SOURCE human.
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 359)
Dias Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Briones,M.R.,
Nagai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.F.,
Goldman,G.H., Carvalho,A.F., Matsukuma,A., Baia,G.S., Simpson,D.H.,
Brunstein,A., deOliveira,P.S., Bucher,P., Jongeneel,C.V., O'Hare
,M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and
Simpson,A.J.
Shotgun sequencing of the human transcriptome with ORF expressed
sequence tags
Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
20202663
Contact: Simpson A.J.G.
Laboratory of Cancer Genetics
Ludwig Institute for Cancer Research
Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
Brazil
Tel: +55-11-2704922
Fax: +55-11-2707001
Email: asimpson@ludwig.org.br
This sequence was derived from the FAPESP/LICR Human Cancer Genome
Project. This entry can be seen in the following URL
(http://www.ludwig.org.br/scripts/gethtml2.pl?tl=&t2=CM4-HT0652-150
400-143-a07&t3=2000-04-15&t4=1)
Seq primer: puc 18 forward
High quality sequence start: 39
High quality sequence stop: 359.
FEATURES
Location/Qualifiers
1..359
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_lib="HT0652"
/dev_stage="Adult"
/note="Organ: head neck; Vector: puc18; Site_1: SmaI;
Site_2: SmaI; A mini-library was made by cloning products
derived from ORESTES PCR (U.S. Letters Patent application
No. 196,716 - Ludwig Institute for Cancer Research)
profiles into the pUC 18 vector. Reverse transcription of
tissue mRNA and cDNA amplification were performed under
low stringency conditions."
BASE COUNT 92 a 90 c 95 g 82 t
ORIGIN
Query Match 60.5%; Score 124.6; DB 10; Length 359;
Best Local Similarity 81.0%; Pred. No. 2.8e-29;
Matches 145; Conservative 0; Mismatches 34; Indels 0; Gaps 0;
QY 28 ATGCTCAAGAACCCAGTTAAAGGTCCTGTGCTACTAAGCCAGGTTCTTGCTCTATTATCT 87
Db 170 AAGCGCAAGAGCCAGTCAAAGGTCCTGCTCCACTAAGCCTGGCTCTGCCCCATTATCT 229
QY 88 TGATTCGTTGGCGTATGTTAAACCCACCTAACCGTTGTTTGAAGGACACTGATTGTCCAG 147
Db 230 TGATCCGGTGGCGCATGTTGAATCCCCCTAACCGTCTGCTTGAAGATACTGACTGCCAG 289
QY 148 GTATCAAAAAGTGCTGTGAAGGTTCCTCGGGTATGGCTTGTTCGTTCCACAATAATAG 206
Db 290 GAATCAAGAAGTGCTGTGAAGGCTCTTCGGGGATGGCCTGTTTCGTTCCCCAGTGAGAG 348
RESULT 9
BE772161/c
LOCUS
DEFINITION CM4-FT0104-230600-215-c09 FT0104 Homo sapiens cDNA, mRNA sequence.
ACCESSION BE772161
VERSION BE772161.1 GI:10225819

AUTHORS
Dias Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Briones,M.R., Nagai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.F., Goldman,G.H., Carvalho,A.F., Matsukuma,A., Baia,G.S., Simpson,D.H., Brunstein,A., deOliveira,P.S., Bucher,P., Jongeneel,C.V., O'Hare ,M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and Simpson,A.J.

TITLE
Shotgun sequencing of the human transcriptome with ORF expressed sequence tags

JOURNAL
MEDLINE
COMMENT
Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)

Contact: Simpson A.J.G.
Laboratory of Cancer Genetics
Ludwig Institute for Cancer Research
Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP, Brazil
Tel: +55-11-2704922
Fax: +55-11-2707001
Email: asimpson@ludwig.org.br

This sequence was derived from the FAPESP/LICR Human Cancer Genome Project. This entry can be seen in the following URL
(http://www.ludwig.org.br/scripts/gethtml2.pl?tl=&t2=CM4-HT0744-160
600-201-al0&t3=2000-06-16&t4=1)

Seq primer: puc 18 forward
High quality sequence start: 11
High quality sequence stop: 370.

FEATURES
source
Location/Qualifiers
1. .370
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_lib="HT0744"
/dev_stage="Adult"
/note="Organ: head neck; Vector: puc18; Site_1: Smal;
Site 2: Smal; A mini-library was made by cloning products
derived from ORESTES PCR (U.S. Letters Patent application
No. 196,716 - Ludwig Institute for Cancer Research)
profiles into the pUC 18 vector. Reverse transcription of
tissue mRNA and cDNA amplification were performed under
low stringency conditions."
BASE COUNT 85 a 97 c 99 g 89 t

Query Match 60.5%; Score 124.6; DB 12; Length 370;
Best Local Similarity 81.0%; Pred. No. 2.8e-29;
Matches 145; Conservative 0; Mismatches 34; Indels 0; Gaps 0;

Qy 28 ATGCTCAAGAACCCAGTTAAAGGTCCTGTGCTACTAAGCCAGGTTCTTGTCTCTATTATCT 87
Db 242 AAGCGCAAGAGCCAGTCAAAGGTCAGTCTCCACTAAGCCTGGCTCCTGCCCATATCT 183
Qy 88 TGATTCGTTGCGCTATGTTAAACCCACCTAACCGTTGTTTGAAGGACACTGATTGTCCAG 147
Db 182 TGATCCGGTGCGCCATGTTGAATCCCCCTAACCGTGCTTGAAGATACTGACTGCCAG 123
Qy 148 GTATCAAAAAGTGTGAAGGTTCTCCTGCGGTATGGCTTGTTCGTTCCACAATAATAG 206
Db 122 GAATCAAGAAGTGCTGTGAAGGCTCTTGGGGATGGCCTGTTTCGTTCCCCAGTGAGAG 64

RESULT 12
BE715626
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 370)
Dias Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Briones,M.R.,
Nagai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.F.,
Goldman,G.H., Carvalho,A.F., Matsukuma,A., Baia,G.S., Simpson,D.H.,

REFERENCE
AUTHORS
Brunstein,A., deOliveira,P.S., Bucher,P., Jongeneel,C.V., O'Hare ,M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and Simpson,A.J.

TITLE
JOURNAL
MEDLINE
COMMENT
Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)

Contact: Simpson A.J.G.
Laboratory of Cancer Genetics
Ludwig Institute for Cancer Research
Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP, Brazil
Tel: +55-11-2704922
Fax: +55-11-2707001
Email: asimpson@ludwig.org.br

This sequence was derived from the FAPESP/LICR Human Cancer Genome Project. This entry can be seen in the following URL
(http://www.ludwig.org.br/scripts/gethtml2.pl?tl=&t2=CM4-HT0744-160
600-201-h09&t3=2000-06-16&t4=1)

Seq primer: puc 18 forward
High quality sequence start: 13
High quality sequence stop: 370.

FEATURES
source
Location/Qualifiers
1. .370
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_lib="HT0744"
/dev_stage="Adult"
/note="Organ: head neck; Vector: puc18; Site_1: Smal;
Site_2: Smal; A mini-library was made by cloning products
derived from ORESTES PCR (U.S. Letters Patent application
No. 196,716 - Ludwig Institute for Cancer Research)
profiles into the pUC 18 vector. Reverse transcription of
tissue mRNA and cDNA amplification were performed under
low stringency conditions."
BASE COUNT 85 a 110 c 87 g 88 t

Query Match 60.5%; Score 124.6; DB 12; Length 370;
Best Local Similarity 81.0%; Pred. No. 2.8e-29;
Matches 145; Conservative 0; Mismatches 34; Indels 0; Gaps 0;

Qy 28 ATGCTCAAGAACCCAGTTAAAGGTCCTGTGCTACTAAGCCAGGTTCTTGTCTATTATCT 87
Db 95 AAGCGCAAGAGCCAGTCAAAGGTCAGTCTCCACTAAGCCTGGCTCCTGCCCATATCT 154
Qy 88 TGATTCGTTGCGCTATGTTAAACCCACCTAACCGTTGTTTGAAGGACACTGATTGTCCAG 147
Db 155 TGATCCGGTGCGCCATGTTGAATCCCCCTAACCGTGCTTGAAGATACTGACTGCCAG 214
Qy 148 GTATCAAAAAGTGTGAAGGTTCTCCTGCGGTATGGCTTGTTCGTTCCACAATAATAG 206
Db 215 GAATCAAGAAGTGCTGTGAAGGCTCTTGGGGATGGCCTGTTTCGTTCCCCAGTGAGAG 273

RESULT 13
BE772160/c
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 370)
Dias Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Briones,M.R.,
Nagai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.F.,
Goldman,G.H., Carvalho,A.F., Matsukuma,A., Baia,G.S., Simpson,D.H.,
Brunstein,A., deOliveira,P.S., Bucher,P., Jongeneel,C.V., O'Hare ,M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and Simpson,A.J.

REFERENCE
AUTHORS
Brunstein,A., deOliveira,P.S., Bucher,P., Jongeneel,C.V., O'Hare ,M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and Simpson,A.J.

Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP, Brazil
Tel: +55-11-2704922
Fax: +55-11-2707001
Email: asimpson@ludwig.org.br
This sequence was derived from the FAPESP/LICR Human Cancer Genome Project. This entry can be seen in the following URL
(http://www.ludwig.org.br/scripts/gethtml2.pl?t1=&t2=CM4-FT0104-230
600-215-d05&t3=2000-06-23&t4=1)
Seq primer: puc 18 forward
High quality sequence start: 8
High quality sequence stop: 371.
High quality sequence stop: 371.

FEATURES
source

1. .371
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_lib="FT0104"
/dev_stage="Adult"
/note="Organ: prostate tumor; Vector: puc18; Site_1: SmaI;
Site 2: SmaI; A mini-library was made by cloning products
derived from ORESTES PCR (U.S. Letters Patent application
No. 196,716 - Ludwig Institute for Cancer Research)
profiles into the pUC 18 vector. Reverse transcription of
tissue mRNA and cDNA amplification were performed under
low stringency conditions."

BASE COUNT 84 a 96 c 101 g 90 t

ORIGIN

Query Match 60.5%; Score 124.6; DB 12; Length 371;
Best Local Similarity 81.0%; Pred. No. 2.8e-29;
Matches 145; Conservative 0; Mismatches 34; Indels 0; Gaps 0;

QY 28 ATGCTCAAGAACCCAGTTAAAGGTCCTGTCTACTAAGCCAGGTTCTTGTCCTATTATCT 87

Db 242 AAGCGCAAGAGCCAGTCAAGAGGTCCAGTCTCCACTAAGCCTGGCTCCTGCCCATTAATCT 183

QY 88 TGATTCGTTGGCTATGTTAAACCCACCTAACCGTTGTTTGAAGGACACTGATTGTCCAG 147

Db 182 TGATCCGGTGGCCCATGTTGAATCCCCCTAACCGTGTCTTGAAGATACACTGACTGCCAG 123

QY 148 GTATCAAAAAGTGTGGAAGGTTCCCTGCGGTATGCGTTGTTTCGTTCCACAATAATAG 206

Db 122 GAATCAAGAAGTGTGTGAAGGCTCTTGGGGGATGGCCTGTTTCGTTCCCGAGTGAGAG 64

RESULT 16
AI285400/c

LOCUS AI285400 375 bp mRNA linear EST 23-NOV-1998
DEFINITION qt69f06.x1 NCI CGAP Eso2 Homo sapiens cDNA clone IMAGE:1960547 3'
similar to gb:S58717_rnal ELAFIN PRECURSOR (HUMAN); contains element
THR repetitive element ; mRNA sequence.

ACCESSION AI285400

VERSION AI285400.1 GI:3923633

KEYWORDS EST.

SOURCE human.

ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 375)
NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.

REFERENCE

AUTHORS National Cancer Institute, Cancer Genome Anatomy Project (CGAP),

TITLE Tumor Gene Index

JOURNAL Unpublished (1997)

COMMENT

Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: Nan Hu, M.D., Ph.D., Mark Roth, M.D., Phillip
Taylor, M.D., Michael R. Emmert-Buck, M.D., Ph.D.
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: Greg Lennon, Ph.D.
cDNA Sequencing by: Washington University Genome Sequencing Center
Clone distribution: NCI-CGAP clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
www-bio.llnl.gov/bbrp/image/image.html

Seq primer: -40UP from Gibco
High quality sequence stop: 356.

FEATURES
source

1. .375
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:1960547"
/clone_lib="NCI CGAP Eso2"
/tissue_type="squamous cell carcinoma"
/lab_host="DH10B"
/note="Organ: esophagus; Vector: pCMV-SPORT6; Site_1: SmaI;
; Site 2: NotI; Cloned unidirectionally. Primer: Oligo
dT. Average insert size 1.1 kb. Life Technologies catalog
#: 11502-010"

BASE COUNT 91 a 83 c 125 g 75 t 1 others

ORIGIN

Query Match 60.5%; Score 124.6; DB 9; Length 375;
Best Local Similarity 81.0%; Pred. No. 2.8e-29;
Matches 145; Conservative 0; Mismatches 34; Indels 0; Gaps 0;

QY 28 ATGCTCAAGAACCCAGTTAAAGGTCCTGTGTCTACTAAGCCAGGTTCTTGTCCTATTATCT 87

Db 351 AAGCGCAAGAGCCAGTCAAGAGGTCCAGTCTCCACTAAGCCTGGCTCCTGCCCATTAATCT 292

QY 88 TGATTCGTTGGCTATGTTAAACCCACCTAACCGTTGTTTGAAGGACACTGATTGTCCAG 147

Db 291 TGATCCGGTGGCCCATGTTGAATCCCCCTAACCGTGTCTTGAAGATACACTGACTGCCAG 232

QY 148 GTATCAAAAAGTGTGGAAGGTTCCCTGCGGTATGCGTTGTTTCGTTCCACAATAATAG 206

Db 231 GAATCAAGAAGTGTGTGAAGGCTCTTGGGGGATGGCCTGTTTCGTTCCCGAGTGAGAG 173

RESULT 17

BE715603/c

LOCUS BE715603 379 bp mRNA linear EST 12-SEP-2000

DEFINITION CM4-HT0744-160600-201-a07 HT0744 Homo sapiens cDNA, mRNA sequence.

ACCESSION BE715603

VERSION BE715603.1 GI:10103868

KEYWORDS EST.

SOURCE human.

ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 379)
Dias Neto, E., Garcia Correa, R., Verjovski-Almeida, S., Briones, M.R.,
Nagai, M.A., da Silva, W. Jr., Zago, M.A., Bordin, S., Costa, F.F.,
Goldman, G.H., Carvalho, A.F., Matsukuma, A., Baia, G.S., Simpson, D.H.,
Brunstein, A., deOliveira, P.S., Bucher, P., Jongeneel, C.V., O'Hare
M.J., Soares, F., Brentani, R.R., Reis, L.F., de Souza, S.J. and
Simpson, A.J.

REFERENCE

AUTHORS

TITLE

Shotgun sequencing of the human transcriptome with ORF expressed
sequence tags

JOURNAL Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)

MEDLINE

COMMENT

Contact: Simpson A.J.G.

Laboratory of Cancer Genetics

Ludwig Institute for Cancer Research

Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,

Brazil

Tel: +55-11-2704922

Fax: +55-11-2707001

Email: asimpson@ludwig.org.br

This sequence was derived from the FAPESP/LICR Human Cancer Genome
Project. This entry can be seen in the following URL
(http://www.ludwig.org.br/scripts/gethtml2.pl?t1=&t2=CM4-HT0744-160
600-201-a07&t3=2000-06-16&t4=1)

Seq primer: puc 18 forward

High quality sequence start: 25

High quality sequence stop: 379.

FEATURES

source

1. .379

```
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_lib="HT0744"
/dev_stage="Adult"
/note="Organ: head neck; Vector: puc18; Site 1: SmaI;
Site 2: SmaI; A mini-library was made by cloning products
derived from ORESTES PCR (U.S. Letters Patent application
No. 196,716 - Ludwig Institute for Cancer Research)
profiles into the pUC 18 vector. Reverse transcription of
tissue mRNA and cDNA amplification were performed under
low stringency conditions."
BASE COUNT      86 a   95 c   112 g   86 t
ORIGIN
Query Match      60.5%; Score 124.6; DB 12; Length 379;
Best Local Similarity 81.0%; Pred. No. 2.9e-29;
Matches 145; Conservative 0; Mismatches 34; Indels 0; Gaps 0;

QY 28 ATGCTCAAGAACCCAGTTAAAGTCTCTGTCTACTAAGCCAGGTTCTTGTCCTATTATCT 87
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 262 AAGCGCAAGAGCCAGTCAAGAGTCCAGTCTCCACTAAGCCCTGGCTCCTGCCCAATTATCT 203
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |

QY 88 TGATTCGTTGGCGCTATGTTAAACCCACCTAACCGTTGTTGAAGGACACTGATTGTCCAG 147
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 202 TGATCCGGTGGCCCATGTTGAATCCCTAACCGTGTCTGAAAGATACTGACTGCCAG 143
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |

QY 148 GTATCAAAAGTGTGTGAAGTTCTCGCGGTATGGCTTGTTCCTCCACAATAATAG 206
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 142 GAATCAAGAAAGTGTGTGAAGGCTCTTGGGGATGGCTGTTCGTTCCCGCAGTGAGAG 84
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |

RESULT 18
BE772155/c
LOCUS      BE772155      380 bp      mRNA      linear      EST 20-SEP-2000
DEFINITION CM4-FT0104-230600-215-b05 FT0104 Homo sapiens cDNA, mRNA sequence.
ACCESSION  BE772155
VERSION    BE772155.1 GI:10225813
KEYWORDS   EST.
SOURCE     human.
ORGANISM   Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 380)
Dias Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Briones,M.R.,
Nagai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.F.,
Goldman,G.H., Carvalho,A.F., Matsukuma,A., Baia,G.S., Simpson,D.H.,
Brunstein,A., deOliveira,P.S., Bucher,P., Jongeneel,C.V., O'Hare
,M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and
Simpson,A.J.
Shotgun sequencing of the human transcriptome with ORF expressed
sequence tags
Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
20202663
Contact: Simpson A.J.G.
Laboratory of Cancer Genetics
Ludwig Institute for Cancer Research
Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
Brazil
Tel: +55-11-2704922
Fax: +55-11-2707001
Email: asimpson@ludwig.org.br
This sequence was derived from the FAPESP/LICR Human Cancer Genome
Project. This entry can be seen in the following URL
(http://www.ludwig.org.br/scripts/gethtml2.pl?tl=&t2=CM4-FT0104-230
600-215-b05&t3=2000-06-23&t4=1)
Seq primer: puc 18 forward
High quality sequence start: 10
High quality sequence stop: 365.
High quality sequence stop: 365.
Location/Qualifiers
1. .380
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_lib="FT0104"

FEATURES
source
```

```
/dev_stage="Adult"
/note="Organ: prostate tumor; Vector: puc18; Site 1: SmaI;
Site 2: SmaI; A mini-library was made by cloning products
derived from ORESTES PCR (U.S. Letters Patent application
No. 196,716 - Ludwig Institute for Cancer Research)
profiles into the pUC 18 vector. Reverse transcription of
tissue mRNA and cDNA amplification were performed under
low stringency conditions."
BASE COUNT      89 a   99 c   100 g   92 t
ORIGIN
Query Match      60.5%; Score 124.6; DB 12; Length 380;
Best Local Similarity 81.0%; Pred. No. 2.9e-29;
Matches 145; Conservative 0; Mismatches 34; Indels 0; Gaps 0;

QY 28 ATGCTCAAGAACCCAGTTAAAGTCTCTGTCTACTAAGCCAGGTTCTTGTCCTATTATCT 87
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 249 AAGCGCAAGAGCCAGTCAAGAGTCCAGTCTCCACTAAGCCCTGGCTCCTGCCCAATTATCT 190
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |

QY 88 TGATTCGTTGGCGCTATGTTAAACCCACCTAACCGTTGTTGAAGGACACTGATTGTCCAG 147
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 189 TGATCCGGTGGCCCATGTTGAATCCCTAACCGTGTCTGAAAGATACTGACTGCCAG 130
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |

QY 148 GTATCAAAAGTGTGTGAAGTTCTCGCGGTATGGCTTGTTCCTCCACAATAATAG 206
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 129 GAATCAAGAAAGTGTGTGAAGGCTCTTGGGGATGGCTGTTCGTTCCCGCAGTGAGAG 71
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |

RESULT 19
BE772171/c
LOCUS      BE772171      380 bp      mRNA      linear      EST 20-SEP-2000
DEFINITION CM4-FT0104-230600-215-g07 FT0104 Homo sapiens cDNA, mRNA sequence.
ACCESSION  BE772171
VERSION    BE772171.1 GI:10225829
KEYWORDS   EST.
SOURCE     human.
ORGANISM   Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 380)
Dias Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Briones,M.R.,
Nagai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.F.,
Goldman,G.H., Carvalho,A.F., Matsukuma,A., Baia,G.S., Simpson,D.H.,
Brunstein,A., deOliveira,P.S., Bucher,P., Jongeneel,C.V., O'Hare
,M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and
Simpson,A.J.
Shotgun sequencing of the human transcriptome with ORF expressed
sequence tags
Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
20202663
Contact: Simpson A.J.G.
Laboratory of Cancer Genetics
Ludwig Institute for Cancer Research
Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
Brazil
Tel: +55-11-2704922
Fax: +55-11-2707001
Email: asimpson@ludwig.org.br
This sequence was derived from the FAPESP/LICR Human Cancer Genome
Project. This entry can be seen in the following URL
(http://www.ludwig.org.br/scripts/gethtml2.pl?tl=&t2=CM4-FT0104-230
600-215-g07&t3=2000-06-23&t4=1)
Seq primer: puc 18 forward
High quality sequence start: 23
High quality sequence stop: 380.
High quality sequence stop: 380.
Location/Qualifiers
1. .380
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_lib="FT0104"
/dev_stage="Adult"
/note="Organ: prostate tumor; Vector: puc18; Site 1: SmaI;
Site 2: SmaI; A mini-library was made by cloning products
```


Db 234 GAATCAAGAAGTGTGTAAGGCTCTTGGGGATGGCTGTTTCGTTCCCGAGTGAGAG 292

RESULT 24

AI393896/c

LOCUS

DEFINITION

AI393896

cg05e09.x1 NCI CGAP CLL1 Homo sapiens cDNA clone IMAGE:2107912 3'

similar to gb:558717_rnal ELAFIN PRECURSOR (HUMAN); contains element

THR repetitive element ;, mRNA sequence.

AI393896

AI393896.1 GI:4223443

EST.

human.

ORGANISM

Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

1 (bases 1 to 390)

NCI-CGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.

National Cancer Institute, Cancer Genome Anatomy Project (CGAP),

Tumor Gene Index

Unpublished (1997)

Contact: Robert Strausberg, Ph.D.

Email: cgapbs-r@mail.nih.gov

Tissue Procurement: Ash Alizadeh, John Byrd, M.D., Mike Grever,

M.D., Louis M. Staudt, M.D., Ph.D.

cDNA Library Preparation: M. Bento Soares, Ph.D.

cDNA Library Arrayed by: Greg Lennon, Ph.D.

DNA Sequencing by: Washington University Genome Sequencing Center

Clone distribution: NCI-CGAP clone distribution information can be

found through the I.M.A.G.E. Consortium/LLNL at:

www-bio.llnl.gov/bbrp/image/image.html

Insert Length: 802 Std Error: 0.00

Seq primer: -40UP from Gibco

High quality sequence stop: 381.

Location/Qualifiers

1..390

/organism="Homo sapiens"

/db_xref="taxon:9606"

/clone="IMAGE:2107912"

/clone_lib="NCI-CGAP_CLL1"

/tissue_type="B-cell, chronic lymphocytic leukemia"

/lab_host="DH10B"

/note="Vector: pT7T3D-Pac (Pharmacia) with a modified

polylinker; Site_1: Not I; Site_2: Eco RI; 1st strand cDNA

was primed with a Not I - oligo(dT) primer [5'

TGTTACCAATCTGAAGTGGAGCGCGCATGCTTTTTTTTTTTTTTTTTTTT

T 3']; double-stranded cDNA was ligated to Eco RI

adaptors (Pharmacia), digested with Not I and cloned into

the Not I and Eco RI sites of the modified pT7T3 vector.

Library is normalized, and was constructed by Bento

Soares and M.Fatima Bonaldo."

BASE COUNT 95 a 86 c 126 g 83 t

ORIGIN

Query Match 60.5%; Score 124.6; DB 9; Length 390;

Best Local Similarity 81.0%; Pred. No. 2.9e-29;

Matches 145; Conservative 0; Mismatches 34; Indels 0; Gaps 0;

Qy 28 ATGCTCAAGAACCAGTTAAAGTCTGTGCTACTAAGCCAGGTTCTTGTCCTATTATCT 87

Db 352 AAGCGCAAGAGCCAGTCAAGAGTCCAGTCTCCACTAAGCTGGCTCCTGCCCATATCT 293

Qy 88 TGATTGCTGCGCTATGTTAAACCCACCTAACCGTTGTTGAAGGACACTGATGTCCAG 147

Db 292 TGATCCGGTGGCCATGTTGAATCCCTTAACCGTCTGCTTGAAGATATACTGACTGCCAG 233

Qy 148 GTATCAAAAAGTGTGTAAGGTTCTCGGGTATGGCTGTTTCGTTCCACAATAATAG 206

Db 232 GAATCAAGAAGTGTGTGAAGGCTCTTGGGGATGGCCTGTTTCGTTCCCGAGTGAGAG 174

RESULT 25

AI989737/c

LOCUS

DEFINITION

AI989737

wu21e08.x1 Soares Dieckgraefe colon NHCD Homo sapiens cDNA clone

IMAGE:2520710 3' similar to gb:558717_rnal ELAFIN PRECURSOR (HUMAN

);, mRNA sequence.

AI989737.1 GI:5836618

EST.

human.

ORGANISM

Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

1 (bases 1 to 390)

NCI-CGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.

National Cancer Institute, Cancer Genome Anatomy Project (CGAP),

Tumor Gene Index

Unpublished (1997)

Contact: Robert Strausberg, Ph.D.

Email: cgapbs-r@mail.nih.gov

This clone is available royalty-free through LLNL; contact the

IMAGE Consortium (info@image.llnl.gov) for further information.

Seq primer: -40UP from Gibco.

Location/Qualifiers

1..390

/organism="Homo sapiens"

/db_xref="taxon:9606"

/clone="IMAGE:2520710"

/clone_lib="Soares_Dieckgraefe_colon_NHCD"

/tissue_type="colonic mucosa from 3 patients with Crohn's

disease"

/lab_host="DH10B (phage-resistant)"

/note="Organ: colon; Vector: pT7T3D-Pac (Pharmacia) with a

modified polylinker; Site_1: Not I; Site_2: Eco RI; 1st

strand cDNA was primed with a Not I - oligo(dT) primer [5'

TGTTACCAATCTGAAGTGGAGCGCGCATGCTTTTTTTTTTTTTTTTTTTT

double-stranded cDNA was ligated to Eco RI adaptors

(Pharmacia), digested with Not I and cloned into the Not I

and Eco RI sites of the modified pT7T3 vector. Library

went through one round of normalization. Tissue samples

provided by Dr. Brian Dieckgraefe (Washington University,

dieck@im.wustl.edu); colonic mucosa represents a range of

disease involvement from moderate to severe Crohn's

disease; samples include both perforating (fistulas) and

non-perforating samples. Library constructed by Bento

Soares and M. Fatima Bonaldo."

BASE COUNT 94 a 86 c 126 g 84 t

ORIGIN

Query Match 60.5%; Score 124.6; DB 9; Length 390;

Best Local Similarity 81.0%; Pred. No. 2.9e-29;

Matches 145; Conservative 0; Mismatches 34; Indels 0; Gaps 0;

Qy 28 ATGCTCAAGAACCAGTTAAAGTCTGTGCTACTAAGCCAGGTTCTTGTCCTATTATCT 87

Db 356 AAGCGCAAGAGCCAGTCAAGAGTCCAGTCTCCACTAAGCTGGCTCCTGCCCATATCT 297

Qy 88 TGATTGCTGCGCTATGTTAAACCCACCTAACCGTTGTTGAAGGACACTGATGTCCAG 147

Db 296 TGATCCGGTGGCCATGTTGAATCCCTTAACCGTCTGCTTGAAGATATACTGACTGCCAG 237

Qy 148 GTATCAAAAAGTGTGTAAGGTTCTCGGGTATGGCTGTTTCGTTCCACAATAATAG 206

Db 236 GAATCAAGAAGTGTGTGAAGGCTCTTGGGGATGGCCTGTTTCGTTCCCGAGTGAGAG 178

RESULT 26

BE715611/c

LOCUS

DEFINITION

BE715611

CM4-HT0744-160600-201-c07 HT0744 Homo sapiens cDNA, mRNA sequence.

ACCESSION

BE715611

VERSION

BE715611.1 GI:10103876

KEYWORDS

EST.

SOURCE

human.

Goldman,G.H., Carvalho,A.F., Matsukuma,A., Baia,G.S., Simpson,D.H., Brunstein,A., deOliveira,P.S., Bucher,P., Jongeneel,C.V., O'Hare ,M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and Simpson,A.J.

Shotgun sequencing of the human transcriptome with ORF expressed sequence tags

Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)

JOURNAL MEDLINE COMMENT

Contact: Simpson A.J.G.

Laboratory of Cancer Genetics

Ludwig Institute for Cancer Research

Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP, Brazil

Tel: +55-11-2704922

Fax: +55-11-2707001

Email: asimpson@ludwig.org.br

This sequence was derived from the FAPESP/LICR Human Cancer Genome Project. This entry can be seen in the following URL

(<http://www.ludwig.org.br/scripts/gethtml2.pl?t1=&t2=CM4-FT0104-230600-215-al2&t3=2000-06-23&t4=1>)

Seq primer: puc 18 forward

High quality sequence start: 4

High quality sequence stop: 397.

Location/Qualifiers

1..397

/organism="Homo sapiens"

/db_xref="taxon:9606"

/clone_lib="FT0104"

/dev_stage="Adult"

Note="Organ: prostate tumor; Vector: puc18; Site_1: SmaI; Site_2: SmaI; A mini-library was made by cloning products derived from ORESTES PCR (U.S. Letters Patent application No. 196,716 - Ludwig Institute for Cancer Research)

profiles into the pUC 18 vector. Reverse transcription of tissue mRNA and cDNA amplification were performed under low stringency conditions."

90 a 118 c 95 g 94 t

BASE COUNT

ORIGIN

Query Match 60.5%; Score 124.6; DB 12; Length 397;

Best Local Similarity 81.0%; Pred. No. 2.9e-29;

Matches 145; Conservative 0; Mismatches 34; Indels 0; Gaps 0;

QY 28 ATGCTCAAGAACCAGTTAAAGGTCCTGTGTCTACTAAGCCAGGTTCTTGTCCTATTATCT 87

Db 122 AAGCGCAAGAGCCAGTCAAGAGTCCAGTCTCCACTAAGCTGGCTCTGCCCATTTATCT 181

QY 88 TGATTCGTTGGCTATGTTAAACCCACCTAACCGTTGTTGAAGGACACTGATTGTCCAG 147

Db 182 TGATCCGGTGGCCCATGTTGAATCCCCCTAACCGCTGCTGAAAGATACTAGTGTCCAG 241

QY 148 GTATCAAAAAGTCTGTGAAGGTTCTCGGATGCTGTTTCGTTCCACAATAATAG 206

Db 242 GAATCAAGAAGTCTGTGAAGGCTCTTGGGGATGGCCTGTTTCGTTCCCCAGTGAGAG 300

RESULT 29

AA586974/c

LOCUS

DEFINITION

AA586974

nn69d08.s1 NCI_CGAP_Lar1 Homo sapiens cDNA clone IMAGE:1089135 3'

similar to gb:S58717.rna1 ELAFIN PRECURSOR (HUMAN);contains element

THR repetitive element ;, mRNA sequence.

AA586974

AA586974.1 GI:2397788

EST.

human.

ORGANISM

Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Euthera; Primates; Catarrhini; Hominidae; Homo.

1 (bases 1 to 399)

REFERENCE

AUTHORS

NCI-CGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.

National Cancer Institute, Cancer Genome Anatomy Project (CGAP),

Tumor Gene Index

JOURNAL COMMENT

Unpublished (1997)

Contact: Robert Strausberg, Ph.D.

Email: cgabbs-r@mail.nih.gov

Tissue Procurement: L. Jeffrey Medeiros, M.D., Michael R.

Emmert-Buck, M.D., Ph.D.

cDNA Library Preparation: Stratagene, Inc.

cDNA Library Arrayed by: Greg Lennon, Ph.D.

DNA Sequencing by: Washington University Genome Sequencing Center

Clone distribution: NCI-CGAP clone distribution information can be

found through the I.M.A.G.E. Consortium/LLNL at:

www-bio.llnl.gov/bbrp/image/image.html

Insert length: 523 Std Error: 0.00

Seq primer: -40ml3 fwd ET from Amersham.

Location/Qualifiers

1..399

/organism="Homo sapiens"

/db_xref="taxon:9606"

/clone="IMAGE:1089135"

/clone_lib="NCI_CGAP_Lar1"

/tissue_type="larynx"

/lab_host="SOLR (kanamycin resistant)"

/note="Organ: larynx; Vector: Bluescript SK-; Site_1:

EcoRI; Site_2: XhoI; Cloned unidirectionally. Primer:

Oligo dT. Larynx. 5' adaptor sequence: 5' GAATCGGCACGAG

3' 3' adaptor sequence: 5' CTCGAGTTTTTTTTTTTTTTT 3'

Average insert size: 0.9 kb."

97 a 88 c 129 g 85 t

BASE COUNT

ORIGIN

Query Match 60.5%; Score 124.6; DB 9; Length 399;

Best Local Similarity 81.0%; Pred. No. 2.9e-29;

Matches 145; Conservative 0; Mismatches 34; Indels 0; Gaps 0;

QY 28 ATGCTCAAGAACCAGTTAAAGGTCCTGTGTCTACTAAGCCAGGTTCTTGTCCTATTATCT 87

Db 351 AAGCGCAAGAGCCAGTCAAGAGTCCAGTCTCCACTAAGCTGGCTCTGCCCATTTATCT 292

QY 88 TGATTCGTTGGCTATGTTAAACCCACCTAACCGTTGTTGAAGGACACTGATTGTCCAG 147

Db 291 TGATCCGGTGGCCCATGTTGAATCCCCCTAACCGCTGCTGAAAGATACTAGTGTCCAG 232

QY 148 GTATCAAAAAGTCTGTGAAGGTTCTCGGATGCTGTTTCGTTCCACAATAATAG 206

Db 231 GAATCAAGAAGTCTGTGAAGGCTCTTGGGGATGGCCTGTTTCGTTCCCCAGTGAGAG 173

RESULT 30

BE715592

LOCUS

DEFINITION

CM4-HT0744-160600-201-b10 HT0744 Homo sapiens cDNA, mRNA sequence.

ACCESSION

BE715592

VERSION

BE715592.1 GI:10103857

KEYWORDS

EST.

SOURCE

human.

ORGANISM

Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Euthera; Primates; Catarrhini; Hominidae; Homo.

1 (bases 1 to 400)

REFERENCE

AUTHORS

Dias Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Briones,M.R.,

Nagai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.F.,

Goldman,G.H., Carvalho,A.F., Matsukuma,A., Baia,G.S., Simpson,D.H.,

Brunstein,A., deOliveira,P.S., Bucher,P., Jongeneel,C.V., O'Hare

,M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and

Simpson,A.J.

Shotgun sequencing of the human transcriptome with ORF expressed

sequence tags

Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)

20202663

COMMENT

Contact: Simpson A.J.G.

Laboratory of Cancer Genetics

Ludwig Institute for Cancer Research

Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,

Brazil

Tel: +55-11-2704922
Fax: +55-11-2707001
Email: asimpson@ludwig.org.br
This sequence was derived from the FAPESP/LICR Human Cancer Genome Project. This entry can be seen in the following URL
(http://www.ludwig.org.br/scripts/gethtml2.pl?t1=et2=CM4-HT0744-160
600-201-b10&t3=2000-06-16&t4=1)
Seq primer: puc 18 forward
High quality sequence start: 7
High quality sequence stop: 400.

FEATURES
source

1. .400
/location/Qualifiers
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_lib="HT0744"
/dev_stage="Adult"
/note="Organ: head neck; Vector: puc18; Site 1: SmaI;
Site 2: SmaI; A mini-library was made by cloning products
derived from ORESTES PCR (U.S. Letters Patent application
No. 196,716 - Ludwig Institute for Cancer Research)
profiles into the pUC 18 vector. Reverse transcription of
tissue mRNA and cDNA amplification were performed under
low stringency conditions."

BASE COUNT 90 a 115 c 100 g 95 t
ORIGIN

Query Match 60.5%; Score 124.6; DB 12; Length 400;
Best Local Similarity 81.0%; Pred. No. 2.9e-29;
Matches 145; Conservative 0; Mismatches 34; Indels 0; Gaps 0;

QY 28 ATGCTCAGAACACAGTTAAAGGTCCTGTCTACTAAGCCAGGTTCTTGTCCCTATTATCT 87

Db 121 AAGCGCAGAGCCAGTCAAGGTCAGTCTCCACTAAGCCTGGCTCCTGCCCATATCT 180

QY 88 TGATTCGTTGCGCTATGTTAAACCCACCTAACCGTTGTTGAAGGACACTGATTGTCAG 147

Db 181 TGATCCGGTGGCCCATGTTGAATCCCTTAACCGCTGCTTGAAGATACTGACTGCCAG 240

QY 148 GTATCAAAAAGTCTGTGAAGTTCTTCCGGTATGGCTTGTTCGTTCCACAATAATAG 206

Db 241 GAATCAAGAAGTGTGTGAAGGCTCTTGGGGATGGCCCTGTTTCGTTCCCGAGTGGAG 299

RESULT 31
BE715609/c
LOCUS BE715609 403 bp mRNA linear EST 12-SEP-2000
DEFINITION CM4-HT0744-160600-201-c04 HT0744 Homo sapiens cDNA, mRNA sequence.
ACCESSION BE715609
VERSION BE715609.1 GI:10103874
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens

REFERENCE 1 (bases 1 to 403)
AUTHORS Dias Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Briones,M.R., Nagai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.F., Goldman,G.H., Carvalho,A.F., Matsukuma,A., Baia,G.S., Simpson,D.H., Brunstein,A., deOliveira,P.S., Bucher,P., Jongeneel,C.V., O'Hare ,M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and Simpson,A.J.J.
TITLE Shotgun sequencing of the human transcriptome with ORF expressed sequence tags
JOURNAL Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
MEDLINE 20202663
COMMENT Contact: Simpson A.J.G.
Laboratory of Cancer Genetics
Ludwig Institute for Cancer Research
Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP, Brazil
Tel: +55-11-2704922
Fax: +55-11-2707001
Email: asimpson@ludwig.org.br

This sequence was derived from the FAPESP/LICR Human Cancer Genome Project. This entry can be seen in the following URL
(http://www.ludwig.org.br/scripts/gethtml2.pl?t1=et2=CM4-HT0744-160
600-201-c04&t3=2000-06-16&t4=1)
Seq primer: puc 18 forward
High quality sequence start: 12
High quality sequence stop: 403.

FEATURES
source

1. .403
/location/Qualifiers
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_lib="HT0744"
/dev_stage="Adult"
/note="Organ: head neck; Vector: puc18; Site 1: SmaI;
Site 2: SmaI; A mini-library was made by cloning products
derived from ORESTES PCR (U.S. Letters Patent application
No. 196,716 - Ludwig Institute for Cancer Research)
profiles into the pUC 18 vector. Reverse transcription of
tissue mRNA and cDNA amplification were performed under
low stringency conditions."

BASE COUNT 96 a 98 c 113 g 96 t
ORIGIN

Query Match 60.5%; Score 124.6; DB 12; Length 403;
Best Local Similarity 81.0%; Pred. No. 2.9e-29;
Matches 145; Conservative 0; Mismatches 34; Indels 0; Gaps 0;

QY 28 ATGCTCAGAACACAGTTAAAGGTCCTGTCTACTAAGCCAGGTTCTTGTCCCTATTATCT 87

Db 274 AAGCGCAGAGCCAGTCAAGGTCAGTCTCCACTAAGCCTGGCTCCTGCCCATATCT 215

QY 88 TGATTCGTTGCGCTATGTTAAACCCACCTAACCGTTGTTGAAGGACACTGATTGTCAG 147

Db 214 TGATCCGGTGGCCCATGTTGAATCCCTTAACCGCTGCTTGAAGATACTGACTGCCAG 155

QY 148 GTATCAAAAAGTGTGTGAAGTTCTTCCGGTATGGCTTGTTCGTTCCACAATAATAG 206

Db 154 GAATCAAGAAGTGTGTGAAGGCTCTTGGGGATGGCTGTTTCGTTCCCGAGTGGAG 96

RESULT 32
AI874186/c

LOCUS AI874186 405 bp mRNA linear EST 07-MAR-2000
DEFINITION wm50a02.x1 NCI CGAP Ut4 Homo sapiens cDNA clone IMAGE:2439338 3', similar to gb:S58717_rnal ELAFIN PRECURSOR (HUMAN);, mRNA sequence.
ACCESSION AI874186
VERSION AI874186.1 GI:5548235
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens

REFERENCE 1 (bases 1 to 405)
AUTHORS NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP), Tumor Gene Index
JOURNAL Unpublished (1997)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: Christopher Moskaluk, M.D., Ph.D., Michael R. Emmert-Buck, M.D., Ph.D.

CDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: Greg Lennon, Ph.D.
DNA Sequencing by: Washington University Genome Sequencing Center
Clone distribution: NCI-CGAP clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at:
www-bio.llnl.gov/bbrp/image/image.html
Insert Length: 644 Std Error: 0.00
Seq primer: -40UP from Gibco.

FEATURES
source

1. .405
/organism="Homo sapiens"
/db_xref="taxon:9606"

/clone="IMAGE:2439338"
/clone_lib="NCI_CGAP_Ut4"
/tissue_type="serous papillary carcinoma, high grade, 2
pooled tumors"
/lab_host="DH10B"
/note="Organ: uterus; Vector: pCMV-SPORT6; Site_1: SalI;
Site_2: NotI; Cloned unidirectionally. Primer: Oligo dT.
Average insert size 1.48 kb. Life Technologies catalog #:
11542-016"

BASE COUNT 96 a 84 c 125 g 100 t
ORIGIN

Query Match 60.5%; Score 124.6; DB 9; Length 405;
Best Local Similarity 81.0%; Pred. No. 2.9e-29;
Matches 145; Conservative 0; Mismatches 34; Indels 0; Gaps 0;

QY 28 ATGCTCAAGAACCCAGTTAAAGTCTGTCTACTAAGCCAGGTTCTTGTCTCTATTATCT 87
|||||
Db 377 AAGCGCAAGAGCCAGTCAAGGTCCTCACTAAGCCTGGCTCTGCCCCATTATCT 318
|||||

QY 88 TGATTGCTGGGCTATGTTAAACCCACCTAACCGTTGTTTGAAGGACACTGATTGTCCAG 147
|||||

Db 317 TGATCCGGTGGCCCTGTTGAATCCCCCTAACCGCTGCTTGAAAGATACTGACTGCCCGAG 258
|||||

QY 148 GTATCAAAAAGTCTGTGAAGGTTCTCTGCGGTATGGCTTGTTCCTCCACAATAATAG 206
|||||

Db 257 GAATCAAGAAGTCTGTGAAGGCTCTTGCGGGATGGCCTGTTTCGTTCCCGCAGTGAGAG 199
|||||

RESULT 33
AW137392/c

LOCUS AW137392 406 bp mRNA linear EST 29-OCT-1999

DEFINITION UI-H-B11-acs-b-12-0-UI.s1 NCI_CGAP_Sub3 Homo sapiens cDNA clone
IMAGE:2715238 3', mRNA sequence.

ACCESSION AW137392
VERSION AW137392.1 GI:6141710

KEYWORDS EST.

SOURCE human.

ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE 1 (bases 1 to 406)
NCI-CGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.

AUTHORS National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index

TITLE Unpublished (1997)

JOURNAL Contact: Robert Strausberg, Ph.D.
COMMENT Email: cgapbs-r@mail.nih.gov
The sequence contained an oligo-dT track that was present in the
oligonucleotide that was used to prime the synthesis of first
strand cDNA and therefore this may represent a bonafide poly A
tail. cDNA Library Preparation: M.B. Soares Lab Clone distribution:
NCI-CGAP clone distribution information can be found through the
I.M.A.G.E. Consortium/LLNL at:
www-bio.llnl.gov/bbrp/image/image.html
Seq primer: M13 Forward
POLYA=Yes.

FEATURES source

1..406
Location/Qualifiers
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:2715238"
/clone_lib="NCI_CGAP_Sub3"
/lab_host="DH10B (Life Technologies)"
/note="Vector: pT7T3D-Pac (Pharmacia) with a modified
polylinker; Site_1: Not I; Site_2: Eco RI; The
NCI_CGAP_Sub3 library is a subtracted library derived from
the NCI_CGAP_Sub1 library, which is a subtracted library
derived from BI. BI constitutes a mixture of 21
normalized or subtracted NCI_CGAP libraries: NCI_CGAP_Co4
, NCI_CGAP_Pr22, NCI_CGAP_Pr28, NCI_CGAP_Co10,
NCI_CGAP_Co16, NCI_CGAP_Kid5, NCI_CGAP_Kid12,
NCI_CGAP_Kid3, NCI_CGAP_Kid11, NCI_CGAP_Lym2,

NCI_CGAP_Br2, NCI_CGAP_Co8, NCI_CGAP_CLL1, NCI_CGAP_Lei2,
NCI_CGAP_Brn23, NCI_CGAP_Lu5, NCI_CGAP_Lu24,
NCI_CGAP_Lu19, NCI_CGAP_GC4, NCI_CGAP_GC6,
NCI_CGAP_Brn25. These 21 libraries were pooled and a
single-stranded DNA preparation of the resulting mixture
was used as a tracer in a subtractive hybridization with
a driver whose composition is detailed below:
NCI_CGAP_Kid3 pool 1 LLAM 3334-3337, 3682-3683,
3798-3803 (IMAGE CloneIDs 1322376-1323911, 1456008-1456775
, 1500552-1502855); NCI_CGAP_Kid5 pool 1 LLAM 3338-3342
, 3722-3725, 3776-3778 (IMAGE CloneIDs 1323912-1325831,
1471368-1472903, 1492104-1493255); NCI_CGAP_Lu5 pool 1
LLAM 3575-3582, 3851-3854 (IMAGE CloneIDs 1414920-1417991,
1520904-1522439); NCI_CGAP_GC4 pool 1 LLAM 3164-3167,
3716-3720, 3733-3735 (IMAGE CloneIDs 1257096-1258631,
1469064-1470983, 1475592-1476743); NCI_CGAP_Pr22 pool 1
LLAM 2457-2459, 2758-2759, 3062-3068 (IMAGE CloneIDs
985608-986759, 1101192-1101959, 1217928-1220615);
NCI_CGAP_Co10 pool 1 LLAM 2644-2653, 2871-2872 (IMAGE
CloneIDs 1057416-1061255, 1144584-1145351). Subtraction
was performed as previously described [Bonaldo, Lennon &
Soares (1996)]: Normalization and Subtraction: Two
Approaches To Facilitate Gene Discovery. Genome Research
6, 791-806.

TAG_LIB=NCI_CGAP_Co10
TAG_TISSUE=colon
TAG_SEQ=AAACG"
BASE COUNT 95 a 86 c 126 g 99 t
ORIGIN

Query Match 60.5%; Score 124.6; DB 10; Length 406;
Best Local Similarity 81.0%; Pred. No. 2.9e-29;
Matches 145; Conservative 0; Mismatches 34; Indels 0; Gaps 0;

QY 28 ATGCTCAAGAACCCAGTTAAAGTCTGTCTACTAAGCCAGGTTCTTGTCTCTATTATCT 87
|||||

Db 368 AAGCGCAAGAGCCAGTCAAGGTCCTCACTAAGCCTGGCTCTGCCCCATTATCT 309
|||||

QY 88 TGATTGCTGGGCTATGTTAAACCCACCTAACCGTTGTTTGAAGGACACTGATTGTCCAG 147
|||||

Db 308 TGATCCGGTGGCCCTGTTGAATCCCCCTAACCGCTGCTTGAAAGATACTGACTGCCCGAG 249
|||||

QY 148 GTATCAAAAAGTCTGTGAAGGTTCTCTGCGGTATGGCTTGTTCCTCCACAATAATAG 206
|||||

Db 248 GAATCAAGAAGTCTGTGAAGGCTCTTGCGGGATGGCCTGTTTCGTTCCCGCAGTGAGAG 190
|||||

RESULT 34
BE715604

LOCUS BE715604 406 bp mRNA linear EST 12-SEP-2000

DEFINITION CM4-HT0744-160600-201-a08 HT0744 Homo sapiens cDNA, mRNA sequence.

ACCESSION BE715604

VERSION BE715604.1 GI:10103869

KEYWORDS EST.

SOURCE human.

ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE 1 (bases 1 to 406)
Dias Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Briones,M.R.,
Nagai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.F.,
Goldman,G.H., Carvalho,A.F., Matsukuma,A., Baia,G.S., Simpson,D.H.,
Brunstein,A., deOliveira,P.S., Bucher,P., Jongeneel,C.V., O'Hare
M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and
Simpson,A.J.

TITLE Shotgun sequencing of the human transcriptome with ORF expressed
sequence tags

JOURNAL Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)

MEDLINE 20202663

COMMENT Contact: Simpson A.J.G.
Laboratory of Cancer Genetics
Ludwig Institute for Cancer Research
Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,

Brazil
Tel: +55-11-2704922
Fax: +55-11-2707001
Email: asimpson@ludwig.org.br
This sequence was derived from the FAPESP/LICR Human Cancer Genome Project. This entry can be seen in the following URL
(http://www.ludwig.org.br/scripts/gethtml2.pl?t1=&t2=CM4-HT0744-160
600-201-a08&t3=2000-06-16&t4=1)
Seq primer: puc 18 forward
High quality sequence start: 21
High quality sequence stop: 406.
Location/Qualifiers

FEATURES
source

1. .406
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_lib="HT0744"
/dev_stage="Adult"
/note="Organ: head neck; Vector: puc18; Site 1: SmaI;
Site 2: SmaI; A mini-library was made by cloning products
derived from ORESTES PCR (U.S. Letters Patent application
No. 196,716 - Ludwig Institute for Cancer Research)
profiles into the pUC 18 vector. Reverse transcription of
tissue mRNA and cDNA amplification were performed under
low stringency conditions."
94 a 118 c 97 g 97 t

BASE COUNT
ORIGIN

Query Match 60.5%; Score 124.6; DB 12; Length 406;
Best Local Similarity 81.0%; Pred. No. 2.9e-29;
Matches 145; Conservative 0; Mismatches 34; Indels 0; Gaps 0;

QY 28 ATGCTCAAGAACCCAGTTAAAGGTCCTGTCTACTAAGCCAGGTTCTTGTCTCTATTATCT 87

Db 131 AAGCGCAAGAGCCAGTCAAAGGTCCTCACTAAGCCTGGCTCCTGCCCATTTATCT 190

QY 88 TGATTCGTTGGCTATGTTAAACCCACCTAACCGTTGTTTGAAGGACACTGATTGTCCAG 147

Db 191 TGATCCGTTGGCCATGTTGAATCCCCCTAACCGCTGTTTGAAGATACTGACTGCCAG 250

QY 148 GTATCAAAAAGTCTGTGAAGTTCTTCGGGTATGGCTTGTTCGTTCCACAATAATAG 206

Db 251 GAATCAAGAAGTCTGTGAAGGCTCTTCGGGATGGCTGTTCGTTCCCCAGTGAGAG 309

RESULT 35

AA586943/c

LOCUS

DEFINITION
nn69a01.s1 NCI CGAP Lar1 Homo sapiens cDNA clone IMAGE:1089096 3',
similar to gb:S58717_rnal ELAFIN PRECURSOR (HUMAN); contains element
HGR repetitive element ;, mRNA sequence.

ACCESSION

VERSION

AA586943.1 GI:2397757

KEYWORDS

SOURCE

ORGANISM

Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

1 (bases 1 to 419)
NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
Unpublished (1997)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: L. Jeffrey Medeiros, M.D., Michael R.
Emmert-Buck, M.D., Ph.D.
cDNA Library Preparation: Stratagene, Inc.
cDNA Library Arrayed by: Greg Lennon, Ph.D.
DNA Sequencing by: Washington University Genome Sequencing Center
Clone distribution: NCI-CGAP clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
www-bio.llnl.gov/bbrp/image/image.html
Insert Length: 615 Std Error: 0.00

Seq primer: -40ml3 fwd. ET from Amersham
High quality sequence stop: 335.
Location/Qualifiers

FEATURES
source

1. .419
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:1089096"
/clone_lib="NCI_CGAP_Lar1"
/tissue_type="Larynx"
/lab_host="SOLR (kanamycin resistant)"
/note="Organ: larynx; Vector: Bluescript SK-; Site 1:
EcoRI; Site 2: XhoI; Cloned unidirectionally. Primer:
Oligo dT. Larynx. 5' adaptor sequence: 5' GAATTCGGCAGAG
3' 3' adaptor sequence: 5' CTCGAGTTTTTTTTTTTTTTT 3',
Average insert size: 0.9 kb."
103 a 92 c 135 g 89 t

BASE COUNT
ORIGIN

Query Match 60.5%; Score 124.6; DB 9; Length 419;
Best Local Similarity 81.0%; Pred. No. 3e-29;
Matches 145; Conservative 0; Mismatches 34; Indels 0; Gaps 0

QY 28 ATGCTCAAGAACCCAGTTAAAGGTCCTGTCTACTAAGCCAGGTTCTTGTCTCTATTATCT 87

Db 351 AAGCGCAAGAGCCAGTCAAAGGTCCTCACTAAGCCTGGCTCCTGCCCATTTATCT 292

QY 88 TGATTCGTTGGCTATGTTAAACCCACCTAACCGTTGTTTGAAGGACACTGATTGTCCAG 147

Db 291 TGATCCGTTGGCCATGTTGAATCCCCCTAACCGCTGCTTGAAGATACTGACTGCCAG 232

QY 148 GTATCAAAAAGTCTGTGAAGTTCTTCGGGTATGGCTTGTTCGTTCCACAATAATAG 206

Db 231 GAATCAAGAAGTCTGTGAAGGCTCTTCGGGATGGCTGTTCGTTCCCCAGTGAGAG 173

RESULT 36

AA082097/c

LOCUS

DEFINITION

xb60d08.x1 NCI CGAP Eso2 Homo sapiens cDNA clone IMAGE:2580687 3',
similar to gb:S58717_rnal ELAFIN PRECURSOR (HUMAN); contains element
MER28 MER28 repetitive element ;, mRNA sequence.

ACCESSION

VERSION

AA082097.1 GI:6037249

KEYWORDS

SOURCE

ORGANISM

Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

1 (bases 1 to 422)
NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
Unpublished (1997)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: Nan Hu, M.D., Ph.D., Mark Roth, M.D., Phillip
Taylor, M.D., Michael R. Emmert-Buck, M.D., Ph.D.
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: Greg Lennon, Ph.D.
DNA Sequencing by: Washington University Genome Sequencing Center
Clone distribution: NCI-CGAP clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
www-bio.llnl.gov/bbrp/image/image.html

Possible reversed clone: polyT not found

Seq primer: -40UP from Gibco

High quality sequence stop: 407.

FEATURES
source

1. .422
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:2580687"
/clone_lib="NCI_CGAP_Eso2"

Db 269 TGATCCGGTGGCCCATGTTGAATCCCCCTAACCCGCTGCTTGAAAGATACTGACTGCCAG 210
QY 148 GTATCAAAAAGTCTGAAGGTTCTCGGGTATGGCTTGTTCGTTCCACAATAATAG 206
Db 209 GAATCAAGAAGTCTGTGAAGGCTCTTGCGGGATGGCCTGTTTCGTTCCCCAGTGAGAG 151

RESULT 41
BF836743
LOCUS
DEFINITION CM2-HT0969-181100-509-b06 HT0969 Homo sapiens cDNA, mRNA EST 13-JAN-2001
ACCESSION BF836743
VERSION BF836743.1 GI:12188790
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 441)
AUTHORS Dias Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Briones,M.R., Nagai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.F., Goldman,G.H., Carvalho,A.F., Matsukuma,A., Baia,G.S., Simpson,D.H., Brunstein,A., deOliveira,P.S., Bucher,P., Jongeneel,C.V., O'Hare ,M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and Simpson,A.J.
TITLE Shotgun sequencing of the human transcriptome with ORF expressed sequence tags
JOURNAL Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
MEDLINE 20202663
COMMENT Contact: Simpson A.J.G.
Laboratory of Cancer Genetics
Ludwig Institute for Cancer Research
Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP, Brazil
Tel: +55-11-2704922
Fax: +55-11-2707001
Email: asimpson@ludwig.org.br
This sequence was derived from the FAPESP/LICR Human Cancer Genome Project. This entry can be seen in the following URL
(http://www.ludwig.org.br/scripts/gethtml2.pl?tl=CM2&t2=CM2-HT0969-181100-509-b06&t3=2000-11-18&t4=1)
Seq primer: puc 18 forward
High quality sequence start: 8
High quality sequence stop: 441.

FEATURES
source
1..441
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_lib="HT0969"
/dev_stage="Adult"
/note="Organ: head neck; Vector: puc18; Site_1: SmaI; Site_2: SmaI; A mini-library was made by cloning products derived from ORESTES PCR (U.S. Letters Patent application No. 196,716 - Ludwig Institute for Cancer Research) profiles into the pUC 18 vector. Reverse transcription of tissue mRNA and cDNA amplification were performed under low stringency conditions."
BASE COUNT 94 a 131 c 109 g 107 t
ORIGIN
Query Match 60.5%; Score 124.6; DB 12; Length 441;
Best Local Similarity 81.0%; Pred. No. 3e-29;
Matches 145; Conservative 0; Mismatches 34; Indels 0; Gaps 0;

QY 28 ATGCTCAAGAACCAGTTAAAGTCTGTGTCTACTAAGCCAGGTTCTTGTCTTATCT 87
Db 146 AAGCGAAGAGCCAGTCAAGGTCAGTCTCCACTAAGCCTGGCTCCTGCCCATATCT 205
QY 88 TGATTCGTTGGCTATGTTAAACCCACCTAACCGTTGTTTGAAGGACACTGATTGTCCAG 147
Db 206 TGATCCGGTGGCCCATGTTGAATCCCCCTAACCGTCTGTTGAAAGATACTGACTGCCAG 265
QY 148 GTATCAAAAAGTCTGTGAAGGTTCTCGGGTATGGCTGTTTCGTTCCACAATAATAG 206

Db 266 GAATCAAGAAGTCTGTGAAGGCTCTTGCGGGATGGCCTGTTTCGTTCCCCAGTGAGAG 324

RESULT 42
AI582329/c
LOCUS
DEFINITION AI582329.1 NCI CGAP Lul9 Homo sapiens cDNA clone IMAGE:2213824 3', similar to gb:S58717_rnal ELAFIN PRECURSOR (HUMAN); contains element THR repetitive element ;, mRNA sequence.
ACCESSION AI582329
VERSION AI582329.1 GI:4568226
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 448)
AUTHORS NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP), Tumor Gene Index
JOURNAL Unpublished (1997)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: Christopher Moskaluk, M.D., Ph.D., Michael R. Emmert-Buck, M.D., Ph.D.
CDNA Library Preparation: M. Bento Soares, Ph.D.
CDNA Library Arrayed by: Greg Lennon, Ph.D.
DNA Sequencing by: Washington University Genome Sequencing Center
Clone distribution: NCI-CGAP clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: www-bio.llnl.gov/bbrp/image/image.html
Insert Length: 598 Std Error: 0.00
Seq primer: -40UP from Gibco
POLYA=No.

FEATURES
Location/Qualifiers
1..448
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:2213824"
/clone_lib="NCI-CGAP_Lul9"
/tissue_type="squamous cell carcinoma, poorly differentiated (4 pooled tumors, including primary and metastatic)"
/dev_stage="adult"
/lab_host="DH10B (phage-resistant)"
/note="Organ: lung; Vector: pT7T3D-Pac (Pharmacia) with a modified polylinker; 1st strand cDNA was prepared from pooled lung tumor tissue, and was then primed with a Not I - oligo(dT) primer. Double-stranded cDNA was ligated to Eco RI adaptors (Pharmacia), digested with Not I and cloned into the Not I and Eco RI sites of the modified pT7T3 vector. Library went through one round of normalization. Library constructed by Bento Soares and M. Fatima Bonaldo."
BASE COUNT 105 a 96 c 138 g 109 t
ORIGIN
Query Match 60.5%; Score 124.6; DB 9; Length 448;
Best Local Similarity 81.0%; Pred. No. 3e-29;
Matches 145; Conservative 0; Mismatches 34; Indels 0; Gaps 0;

QY 28 ATGCTCAAGAACCAGTTAAAGTCTGTGTCTACTAAGCCAGGTTCTTGTCTTATCT 87
Db 366 AAGCGAAGAGCCAGTCAAGGTCAGTCTCCACTAAGCCTGGCTCCTGCCCATATCT 307
QY 88 TGATTCGTTGGCTATGTTAAACCCACCTAACCGTTGTTTGAAGGACACTGATTGTCCAG 147
Db 306 TGATCCGGTGGCCCATGTTGAATCCCCCTAACCGTCTGTTGAAAGATACTGACTGCCAG 247
QY 148 GTATCAAAAAGTCTGTGAAGGTTCTCGGGTATGGCTGTTTCGTTCCACAATAATAG 206
Db 246 GAATCAAGAAGTCTGTGAAGGCTCTTGCGGGATGGCCTGTTTCGTTCCCCAGTGAGAG 188

RESULT 43
AI459240/c

LOCUS AI459240 453 bp mRNA linear EST 14-APR-1999
DEFINITION tk11c09.x1 NCI_CGAP_Lu24 Homo sapiens cDNA clone IMAGE:2150704 3' similar to gb:S58717.rnal ELAFIN PRECURSOR (HUMAN);contains element THR repetitive element ;, mRNA sequence.

ACCESSION AI459240
VERSION AI459240.1 GI:43111819
KEYWORDS EST.
SOURCE human.

ORGANISM

Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 453)

REFERENCE NCI-CGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.

AUTHORS National Cancer Institute, Cancer Genome Anatomy Project (CGAP),

TITLE Tumor Gene Index

JOURNAL Unpublished (1997)

COMMENT Contact: Robert Strausberg, Ph.D.

Email: cgapbs-r@mail.nih.gov

Tissue Procurement: Christopher Moskaluk, M.D., Ph.D., Michael R.

Emmert-Buck, M.D., Ph.D.

cDNA Library Preparation: M. Bento Soares, Ph.D.

cDNA Library Arrayed by: Greg Lennon, Ph.D.

DNA Sequencing by: Washington University Genome Sequencing Center

Clone distribution: NCI-CGAP clone distribution information can be

found through the I.M.A.G.E. Consortium/LLNL at:

www-bio.llnl.gov/bbrp/image/image.html

Insert Length: 554 Std Error: 0.00

Seq primer: -40UP from Gibco.

FEATURES

source

1..453

/organism="Homo sapiens"

/db_xref="taxon:9606"

/clone="IMAGE:2150704"

/clone_lib="NCI_CGAP_Lu24"

/tissue_type="carcinoid"

/lab_host="DH10B"

/note="Organ: lung; Vector: pT7T3D-Pac (Pharmacia) with a modified polylinker; Plasmid DNA from the normalized library NCI_CGAP_Lu5 was prepared, and ss circles were made in vitro. Following HAP purification, this DNA was used as tracer in a subtractive hybridization reaction. The driver was PCR-amplified cDNAs from a pool of 5,000 clones made from the same library (cloneIDs 1414920-1417991 and 1520904-1522439). Subtraction by Bento Soares and M. Fatima Bonaldo."

BASE COUNT 107 a 96 c 137 g 113 t

ORIGIN Soares and M. Fatima Bonaldo.

Query Match 60.5%; Score 124.6; DB 9; Length 453;

Best Local Similarity 81.0%; Pred. No. 3e-29;

Matches 145; Conservative 0; Mismatches 34; Indels 0; Gaps 0;

QY 28 ATGCTCAAGAACCCAGTTAAAGTCTGTGTCTACTAAGCCAGGTTCTTGTCCTATTATCT 87

Db 368 AAGCGCAAGAGCCAGTCAAGGTCAGTCTCCACTAAGCCTGGCTCCTGCCCATATCT 309

QY 88 TGATTCGTTGCGCTATGTTAAACCCACCTAACCGTTGTTTGAAGGACACTGATTGTCCAG 147

Db 308 TGATCCGGTGGCCCATGTTGAATCCCCCTAACCGTCTGTGAAGATACACTGACTGCCAG 249

QY 148 GTATCAAAAAGTGCTGTGAAGGTTCTCGGGTATGGCTTGTTCGTTCCACAATAAG 206

Db 248 GAATCAAGAAGTGCTGTGAAGGCTCTTGGGGATGCGCTGTTTCGTTCCCGAGTGAG 190

RESULT 44

BF824934

LOCUS

DEFINITION ILO-HN0039-101100-500-b06 HN0039 Homo sapiens cDNA, mRNA sequence.

ACCESSION

BF824934

VERSION BF824934.1 GI:12167001

KEYWORDS EST.

SOURCE human.

ORGANISM

Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE 1 (bases 1 to 465)

AUTHORS

Dias Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Briones,M.R., Nagai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.F., Goldman,G.H., Carvalho,A.F., Matsukuma,A., Baia,G.S., Simpson,D.H., Brunstein,A., deOliveira,P.S., Bucher,P., Jongeneel,C.V., O'Hare ,M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and Simpson,A.J.

Shotgun sequencing of the human transcriptome with ORF expressed

sequence tags

JOURNAL Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)

MEDLINE 20202663

COMMENT

Contact: Simpson A.J.G.

Laboratory of Cancer Genetics

Ludwig Institute for Cancer Research

Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP, Brazil

Tel: +55-11-2704922

Fax: +55-11-2707001

Email: asimpson@ludwig.org.br

This sequence was derived from the FAPESP/LICR Human Cancer Genome Project. This entry can be seen in the following URL (<http://www.ludwig.org.br/scripts/gethtml2.pl?tl=ILO&t2=ILO-HN0039-101100-500-b06&t3=2000-11-10&t4=1>)

Seq primer: puc 18 forward

High quality sequence stop: 465.

FEATURES

source

1..465

/organism="Homo sapiens"

/db_xref="taxon:9606"

/clone_lib="HN0039"

/dev_stage="Adult"

/note="Organ: head normal; Vector: puc18; Site_1: SmaI; Site_2: SmaI; A mini-library was made by cloning products derived from ORESTES PCR (U.S. Letters Patent application No. 196,716 - Ludwig Institute for Cancer Research) profiles into the pUC 18 vector. Reverse transcription of tissue mRNA and cDNA amplification were performed under low stringency conditions."

BASE COUNT 98 a 137 c 117 g 113 t

ORIGIN

Query Match 60.5%; Score 124.6; DB 12; Length 465;

Best Local Similarity 81.0%; Pred. No. 3.1e-29;

Matches 145; Conservative 0; Mismatches 34; Indels 0; Gaps 0;

QY 28 ATGCTCAAGAACCCAGTTAAAGTCTGTGTCTACTAAGCCAGGTTCTTGTCCTATTATCT 87

Db 181 AAGCGCAAGAGCCAGTCAAGGTCAGTCTCCACTAAGCCTGGCTCCTGCCCATATCT 240

QY 88 TGATTCGTTGCGCTATGTTAAACCCACCTAACCGTTGTTTGAAGGACACTGATTGTCCAG 147

Db 241 TGATCCGGTGGCCCATGTTGAATCCCCCTAACCGTCTGTGAAGATACACTGACTGCCAG 300

QY 148 GTATCAAAAAGTGCTGTGAAGGTTCTCGGGTATGGCTTGTTCGTTCCACAATAAG 206

Db 301 GAATCAAGAAGTGCTGTGAAGGCTCTTGGGGATGCGCTGTTTCGTTCCCGAGTGAG 359

RESULT 45

BF825207/c

LOCUS

DEFINITION ILO-HN0039-131100-501-d07 HN0039 Homo sapiens cDNA, mRNA sequence.

ACCESSION BF825207

VERSION BF825207.1 GI:12167451

KEYWORDS EST.

SOURCE human.

LOCUS BE181959 312 bp mRNA linear EST 22-JUN-2000
DEFINITION CM1-HT0641-280400-206-d02 HT0641 Homo sapiens cDNA, mRNA sequence.
ACCESSION BE181959
VERSION BE181959.1 GI:8661135
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 312)
AUTHORS Dias Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Briones,M.R.,
Nagai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.F.,
Goldman,G.H., Carvalho,A.F., Matsukuma,A., Baia,G.S., Simpson,D.H.,
Brunstein,A., deOliveira,P.S., Bucher,P., Jongeneel,C.V., O'Hare
,M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and
Simpson,A.J.
TITLE Shotgun sequencing of the human transcriptome with ORF expressed
sequence tags
JOURNAL Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
MEDLINE 20202663
COMMENT Contact: Simpson A.J.G.
Laboratory of Cancer Genetics
Ludwig Institute for Cancer Research
Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
Brazil
Tel: +55-11-2704922
Fax: +55-11-2707001
Email: asimpson@ludwig.org.br
This sequence was derived from the FAPESP/LICR Human Cancer Genome
Project. This entry can be seen in the following URL
(http://www.ludwig.org.br/scripts/gethtml2.pl?tl=&t2=CM1-HT0641-280
400-206-d02&t3=2000-04-28&t4=1)
Seq primer: puc 18 forward
High quality sequence start: 15
High quality sequence stop: 312.
FEATURES
Location/Qualifiers
1..312
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_lib="HT0641"
/dev_stage="Adult"
/note="Organ: head_neck; Vector: puc18; Site_1: SmaI;
Site_2: SmaI; A mini-library was made by cloning products
derived from ORESTES PCR (U.S. Letters Patent application
No. 196,716 - Ludwig Institute for Cancer Research)
profiles into the pUC 18 vector. Reverse transcription of
tissue mRNA and cDNA amplification were performed under
low stringency conditions."

BASE COUNT 70 a 93 c 81 g 68 t
ORIGIN
Query Match 59.7%; Score 123; DB 10; Length 312;
Best Local Similarity 80.4%; Pred. No. 8.7e-29;
Matches 144; Conservative 0; Mismatches 35; Indels 0; Gaps 0;
QY 28 ATGCTCAAGAACCCAGTTAAAGGTCCTGTGTCTACTAAGCCAGGTTCTTGCTCCTATTATCT 87
Db 69 AAGCGCAAGAGCCGGTCAAAGGTCAGTCTCCACTAAGCCTGGCTCCTGCCCATTTATCT 128
QY 88 TGATTCGTTGCGCTATGTTAAACCCACCTAACCGTGTGTTGAAGGACACTGATTGTCCAG 147
Db 129 TGATCCGGTGCGCCATGTTGAATCCCCCTAACCGTGTGTTGAAGATACTGACTGCCAG 188
QY 148 GTATCAAAAAGTCTGTGAAGGTTTCCTGCGGTATGGCTTGTTCGTTCCACAATAATAG 206
Db 189 GAATCAAGAAGTGCTGTGAAGGCTCTTGGGGATGGCTGTTTCGTTCCCGAGTGAGAG 247

RESULT 59
BE715619/c
LOCUS BE715619
DEFINITION CM4-HT0744-160600-201-f01 HT0744 Homo sapiens cDNA, mRNA sequence.
ACCESSION BE715619

VERSION BE715619.1 GI:10103884
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 342)
AUTHORS Dias Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Briones,M.R.,
Nagai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.F.,
Goldman,G.H., Carvalho,A.F., Matsukuma,A., Baia,G.S., Simpson,D.H.,
Brunstein,A., deOliveira,P.S., Bucher,P., Jongeneel,C.V., O'Hare
,M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and
Simpson,A.J.
TITLE Shotgun sequencing of the human transcriptome with ORF expressed
sequence tags
JOURNAL Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
MEDLINE 20202663
COMMENT Contact: Simpson A.J.G.
Laboratory of Cancer Genetics
Ludwig Institute for Cancer Research
Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
Brazil
Tel: +55-11-2704922
Fax: +55-11-2707001
Email: asimpson@ludwig.org.br
This sequence was derived from the FAPESP/LICR Human Cancer Genome
Project. This entry can be seen in the following URL
(http://www.ludwig.org.br/scripts/gethtml2.pl?tl=&t2=CM4-HT0744-160
600-201-f01&t3=2000-06-16&t4=1)
Seq primer: puc 18 forward
High quality sequence start: 15
High quality sequence stop: 342.
FEATURES
Location/Qualifiers
1..342
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_lib="HT0744"
/dev_stage="Adult"
/note="Organ: head_neck; Vector: puc18; Site_1: SmaI;
Site_2: SmaI; A mini-library was made by cloning products
derived from ORESTES PCR (U.S. Letters Patent application
No. 196,716 - Ludwig Institute for Cancer Research)
profiles into the pUC 18 vector. Reverse transcription of
tissue mRNA and cDNA amplification were performed under
low stringency conditions."

BASE COUNT 78 a 87 c 93 g 83 t
ORIGIN
Query Match 59.7%; Score 123; DB 12; Length 342;
Best Local Similarity 80.4%; Pred. No. 9e-29;
Matches 144; Conservative 0; Mismatches 35; Indels 0; Gaps 0;
QY 28 ATGCTCAAGAACCCAGTTAAAGGTCCTGTGTCTACTAAGCCAGGTTCTTGCTCCTATTATCT 87
Db 236 AAGCGCAAGAGCCAGTCAAAGGTCAGTCTCCACTAAGCCTGGCTCCTGCCCATTTATCT 177
QY 88 TGATTCGTTGCGCTATGTTAAACCCACCTAACCGTGTGTTGAAGGACACTGATTGTCCAG 147
Db 176 TGATCCGGTGCGCCACGTTGAATCCCCCTAACCGTGTGTTGAAGATACTGACTGCCAG 117
QY 148 GTATCAAAAAGTGTGTGAAGGTTTCCTGCGGTATGGCTTGTTCGTTCCACAATAATAG 206
Db 116 GAATCAAGAAGTGCTGTGAAGGCTCTTGGGGATGGCTGTTTCGTTCCCGAGTGAGAG 58

RESULT 60
BF094327/c
LOCUS BF094327
DEFINITION CM4-UT0043-090900-309-b10 UT0043 Homo sapiens cDNA, mRNA sequence.
ACCESSION BF094327
VERSION BF094327.1 GI:10900037
KEYWORDS EST.
SOURCE human.

ORGANISM	Homo sapiens Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo. 1 (bases 1 to 345)
REFERENCE	
AUTHORS	Dias Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Briones,M.R., Nagai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.F., Goldman,G.H., Carvalho,A.F., Matsukuma,A., Baia,G.S., Simpson,D.H., Brunstein,A., deOliveira,P.S., Bucher,P., Jongeneel,C.V., O'Hare ,M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and Simpson,A.J.
TITLE	Shotgun sequencing of the human transcriptome with ORF expressed sequence tags
JOURNAL	Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
MEDLINE	20202663
COMMENT	Contact: Simpson A.J.G. Laboratory of Cancer Genetics Ludwig Institute for Cancer Research Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP, Brazil Tel: +55-11-2704922 Fax: +55-11-2707001 Email: asimpson@ludwig.org.br This sequence was derived from the FAPESP/LICR Human Cancer Genome project. This entry can be seen in the following URL (http://www.ludwig.org.br/scripts/gethtml2.pl?t1=&t2=CM4-UT0043-090) 900-309-b10&t3=2000-09-09&t4=1) Seq primer: puc 18 forward High quality sequence start: 16 High quality sequence stop: 345.
FEATURES	Location/Qualifiers source 1..345

```

FEATURES
source
Location/Qualifiers
1. .345
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_lib="UT0043"
/dev_stage="Adult"
/note="Organ: uterus_tumor; Vector: puc18; Site 1: SmaI;
Site 2: SmaI; A mini-library was made by cloning products
derived from ORESTES PCR (U.S. Letters Patent application
No. 196,716 - Ludwig Institute for Cancer Research)
profiles into the pUC 18 vector. Reverse transcription of
tissue mRNA and cDNA amplification were performed under
low stringency conditions."
78 a 86 c 98 g 83 t
BASE COUNT

```

	78 a	86 c	98 g	83 t
BASE COUNT				
ORIGIN				
	Query Match 59.7%; Score 123; DB 12; Length 345;			
	Best Local Similarity 80.4%; Pred. No. 9e-29;			
	Matches 144; Conservative 0; Mismatches 35; Indels 0; Gaps 0;			
QY	28	ATGCTCAAGAACCCAGTTAAAGTTCCTGTGTCTACTAAGCCAGGTTCTTGTCCTATTATCT	87	
Db	237	AAGCGCAAGAGCCAGTCAAAGGTCAGTCTCCACTAAGCCTGGCTCCTGCCCATTAATCT	178	
QY	88	TGATTTCGTTGCGGCTATGTTAAACCCACCTAACCGTTGTTTTGAAGGACACTGATTGTCCAG	147	
Db	177	TGATCCGGTGC GCCATGTTGAATCCCCCTAACCGCTGCTTGAAAGATACTGACTGCCCAG	118	
QY	148	GTATCAAAAAGTGCTGTGAAGGTTCTCTGCGGTATGGCTTGTTTCGTTCCACAATAATAG	206	
Db	117	GAATCAAGAAGTGCTGTGAAGGCTCTTTCGCGGATGGCCCGGTTTCGTTCCCCAGTGAGAG	59	

RESULT 61
BF836201
LOCUS BF836201 374 bp mRNA linear EST 13-JAN-2001
DEFINITION QV3-HT1016-171100-474-c05 HT1016 Homo sapiens cDNA, mRNA sequence.
ACCESSION BF836201
VERSION BF836201.1 GI:12187891
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE	1 (bases 1 to 374)
AUTHORS	Dias Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Briones,M.R., Nagai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.F., Goldman,G.H., Carvalho,A.F., Matsukuma,A., Baia,G.S., Simpson,D.H., Brunstein,A., deOliveira,P.S., Bucher,P., Jongeneel,C.V., O'Hare,M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and Simpson,A.J.
TITLE	Shotgun sequencing of the human transcriptome with ORF expressed sequence tags
JOURNAL	Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
MEDLINE	20202663
COMMENT	Contact: Simpson A.J.G. Laboratory of Cancer Genetics Ludwig Institute for Cancer Research Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP, Brazil Tel: +55-11-2704922 Fax: +55-11-2707001 Email: asimpson@ludwig.org.br This sequence was derived from the FAPESP/LICR Human Cancer Genome Project. This entry can be seen in the following URL (http://www.ludwig.org.br/scripts/gethtml2.pl?t1=QV3&t2=QV3-HT1016-171100-474-c05&t3=2000-11-17&t4=1) Seq primer: puc 18 forward High quality sequence start: 11 High quality sequence stop: 374. Location/Qualifiers 1. .374 /organism="Homo sapiens" /db_xref="taxon:9606" /clone_lib="HT1016"
FEATURES	
source	

FEATURES	source
Location/Qualifiers	
1. .374	
/organism="Homo sapiens"	
/db_xref="taxon:9606"	
/clone_lib="HT1016"	
/dev_stage="Adult"	
/note="Organ: head neck; Vector: puc18; Site 1: SmaI; Site 2: SmaI; A mini-library was made by cloning products derived from ORESTES PCR (U.S. Letters Patent application No. 196,716 - Ludwig Institute for Cancer Research) profiles into the pUC 18 vector. Reverse transcription of tissue mRNA and cDNA amplification were performed under low stringency conditions."	
84 a	98 c 102 g 90 t
BASE COUNT	
ORIGIN	

```

BASE COUNT      84 a      98 c      102 g      90 t
ORIGIN

Query Match          59.7%; Score 123; DB 12; Length 374;
Best Local Similarity 80.4%; Pred. No. 9.3e-29;
Matches 144; Conservative 0; Mismatches 35; Indels 0; Gaps 0;

QY   28 ATGCTCAAGAACCCAGTTAAAGGTCCGTGTCTACTAAGCCAGTTCCTTGTCCTATTATCT 87
     ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db   167 AAGCGCAAGAGCCAGTCAAAGGTCACGTCTCCACTAAGCCTGGCTCCTGCCCATTAATCT 226

QY   88 TGATTCGTTGCCGTATGTTAAACCCACCTAACCGTTGTTTTGAAGGACACTGATTGTCCAG 147
     ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db   227 TGATCCGGTGCGGCATGTTGAATCCCCCTAAGCGCTGCTTGAAAGATACTGACTGCCCCAG 286

QY   148 GTATCAAAAAGTGCTGTGAAGGTTTCCTGCGGHATCGCTTGTTTCGTCCACAATAATAG 206
     ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db   287 GAATCAAGAAGTGCTGTGAAGGCTCTTTGCGGGATGGCTGTGTTTCGTTCGCCAGTGAGAG 345

```

RESULT 62
 BIO21805/c
 LOCUS
 DEFINITION
 ACCESSION
 VERSION
 KEYWORDS
 SOURCE
 ORGANISM
 REFERENCE
 AUTHORS


```
BASE COUNT      100 a      146 c      120 g      121 t
ORIGIN

Query Match      59.7%; Score 123; DB 12; Length 487;
Best Local Similarity 80.4%; Pred. No. 1e-28;
Matches 144; Conservative 0; Mismatches 35; Indels 0; Gaps 0;

QY  28 ATGCTCAAGAACCAAGTTAAAGTCTGTGTCTACTAAGCCAGGTTCTTGTCTCTATTATCT 87
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db  203 AAGCGCAAGAGCCAGTCGAAGGTCCAGTCTCCACTAAGCCTGGCTCTGCCCCATTATCT 262

QY  88 TGATTCGTTGCGCTATGTTAAACCCACCTAACCGTTGTTTGAAGGACACTGATTGTCCAG 147
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db  263 TGATCCGGTGCGCCATGCTGAATCCCCCTTAACCGTGTCTTGAAGATACCTGACTGCCAG 322

QY  148 GTATCAAAAAGTCTGTGAAGGTTCTCTGCGGTATGGCTGTTTCGTTCCACAATAATAG 206
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db  323 GAATCAAGAAGTCTGTGAAGGCTCTTTCGGGATGGCCTGTTTCGTTCCCAAGTGAGAG 381

RESULT 73
AW860590
LOCUS      AW860590
DEFINITION QV0-CT0383-030300-139-d12 CT0383 Homo sapiens cDNA, mRNA sequence.
ACCESSION  AW860590
VERSION    AW860590.1 GI:7956283
KEYWORDS   EST.
SOURCE     human.
ORGANISM   Homo sapiens
REFERENCE  Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS    Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
           1 (bases 1 to 344)
Dias Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Briones,M.R.,
Nagai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.F.,
Goldman,G.H., Carvalho,A.F., Matsukuma,A., Baia,G.S., Simpson,D.H.,
Brunstein,A., deOliveira,P.S., Bucher,P., Jongeneel,C.V., O'Hare
,M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and
Simpson,A.J.
TITLE       Shotgun sequencing of the human transcriptome with ORF expressed
           sequence tags
JOURNAL     Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
MEDLINE    20202663
COMMENT     Contact: Simpson A.J.G.
           Laboratory of Cancer Genetics
           Ludwig Institute for Cancer Research
           Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
           Brazil
           Tel: +55-11-2704922
           Fax: +55-11-2707001
           Email: asimpson@ludwig.org.br
           This sequence was derived from the FAPESP/LICR Human Cancer Genome
           Project. This entry can be seen in the following URL
           (http://www.ludwig.org.br/scripts/gethtml2.pl?tl=&t2=QV0-CT0383-030
           300-139-d12&t3=2000-03-03&t4=1)
           Seq primer: puc 18 forward
           High quality sequence start: 54
           High quality sequence stop: 344.
           Location/Qualifiers
               1..344
                   /organism="Homo sapiens"
                   /db_xref="taxon:9606"
                   /clone_lib="CT0383"
                   /dev_stage="Adult"
                   /note="Organ: colon; Vector: puc18; Site 1: SmaI; Site 2:
                   SmaI; A mini-library was made by cloning products derived
                   from ORESTES PCR (U.S. Letters Patent application No. 196
                   ,716 - Ludwig Institute for Cancer Research) profiles
                   into the pUC 18 vector. Reverse transcription of tissue
                   mRNA and cDNA amplification were performed under low
                   stringency conditions."
BASE COUNT      75 a      102 c      82 g      85 t
ORIGIN

Query Match      59.1%; Score 121.8; DB 9; Length 386;
Best Local Similarity 81.5%; Pred. No. 2.3e-28;
Matches 141; Conservative 0; Mismatches 32; Indels 0; Gaps 0;

QY  34 AAGAACCAAGTTAAAGTCTGTGTCTACTAAGCCAGGTTCTTGTCTCTATCTTCTGATTCT 93
```

```
Query Match      59.1%; Score 121.8; DB 10; Length 344;
Best Local Similarity 77.8%; Pred. No. 2.2e-28;
Matches 147; Conservative 0; Mismatches 42; Indels 0; Gaps 0;

QY  18 ATACCTGCATATGCTCAAGAACCAAGTTAAAGTCTCTGTGTCTACTAAGCCAGGTTCTTGT 77
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db  73 ATAAAGTCTAAAGCGCAAGAGCCAGTCTAAGTCCAGTCTCCACTAAGCTGGCTCCTGC 132

QY  78 CCTATTATCTTGATTCGTTGCGCTATGTTAAACCCACCTAACCGTTGTTTGAAGGACACT 137
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db  133 CCCATTATCTTGATCCGTTGCGCCATGTTTGAATCCCCCTAACCGTCTGTGAAAGATACT 192

QY  138 GATTGTCCAGGTATCAAAAAGTCTGTGAAGGTTCTCTGCGGTATGGCTTGTTCGTTCCA 197
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db  193 GACTGCCCAGGAATCAAGAAGAGCTGTGAAGGCTCTTTCGGGATGGCCTGTTTCGTTCCC 252

QY  198 CAATAATAG 206
    ||| ||| |||
Db  253 CAGTGAGAG 261

RESULT 74
AA586983/c
LOCUS      AA586983
DEFINITION nn69e10.s1 NCI CGAP Lar1 Homo sapiens cDNA clone IMAGE:1089162 3'
           similar to gb:S58717_rnal ELAFIN PRECURSOR (HUMAN); contains element
           THR repetitive element ;, mRNA sequence.
ACCESSION  AA586983
VERSION    AA586983.1 GI:2397797
KEYWORDS   EST.
SOURCE     human.
ORGANISM   Homo sapiens
REFERENCE  Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS    Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
           1 (bases 1 to 386)
NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
JOURNAL     Unpublished (1997)
COMMENT     Contact: Robert Strausberg, Ph.D.
           Email: cgapbs-r@mail.nih.gov
           Tissue Procurement: L. Jeffrey Medeiros, M.D., Michael R.
           Emmert-Buck, M.D., Ph.D.
           cDNA Library Preparation: Stratagene, Inc.
           cDNA Library Arrayed by: Greg Lennon, Ph.D.
           DNA Sequencing by: Washington University Genome Sequencing Center
           Clone distribution: NCI-CGAP clone distribution information can be
           found through the I.M.A.G.E. Consortium/LLNL at:
           www-bio.llnl.gov/bbrp/image/image.html
           Insert Length: 642 Std Error: 0.00
           Seq primer: -40ml3 fwd. ET from Amersham
           High quality sequence stop: 241.
           Location/Qualifiers
               1..386
                   /organism="Homo sapiens"
                   /db_xref="taxon:9606"
                   /clone="IMAGE:1089162"
                   /clone_lib="NCI-CGAP_Lar1"
                   /tissue_type="larynx"
                   /lab_host="SOLR (kanamycin resistant)"
                   /note="Organ: larynx; Vector: Bluescript SK-; Site_1:
                   EcoRI; Site_2: XhoI; Cloned unidirectionally. Primer:
                   Oligo dT. Larynx. 5' adaptor sequence: 5' GAATTCGGCAGG
                   3' 3' adaptor sequence: 5' CTCGAGTTTTTTTTTTTTTTT 3'
                   Average insert size: 0.9 kb."
BASE COUNT      86 a      75 c      121 g      104 t
ORIGIN

Query Match      59.1%; Score 121.8; DB 9; Length 386;
Best Local Similarity 81.5%; Pred. No. 2.3e-28;
Matches 141; Conservative 0; Mismatches 32; Indels 0; Gaps 0;

QY  34 AAGAACCAAGTTAAAGTCTGTGTCTACTAAGCCAGGTTCTTGTCTCTATCTTCTGATTCT 93
```

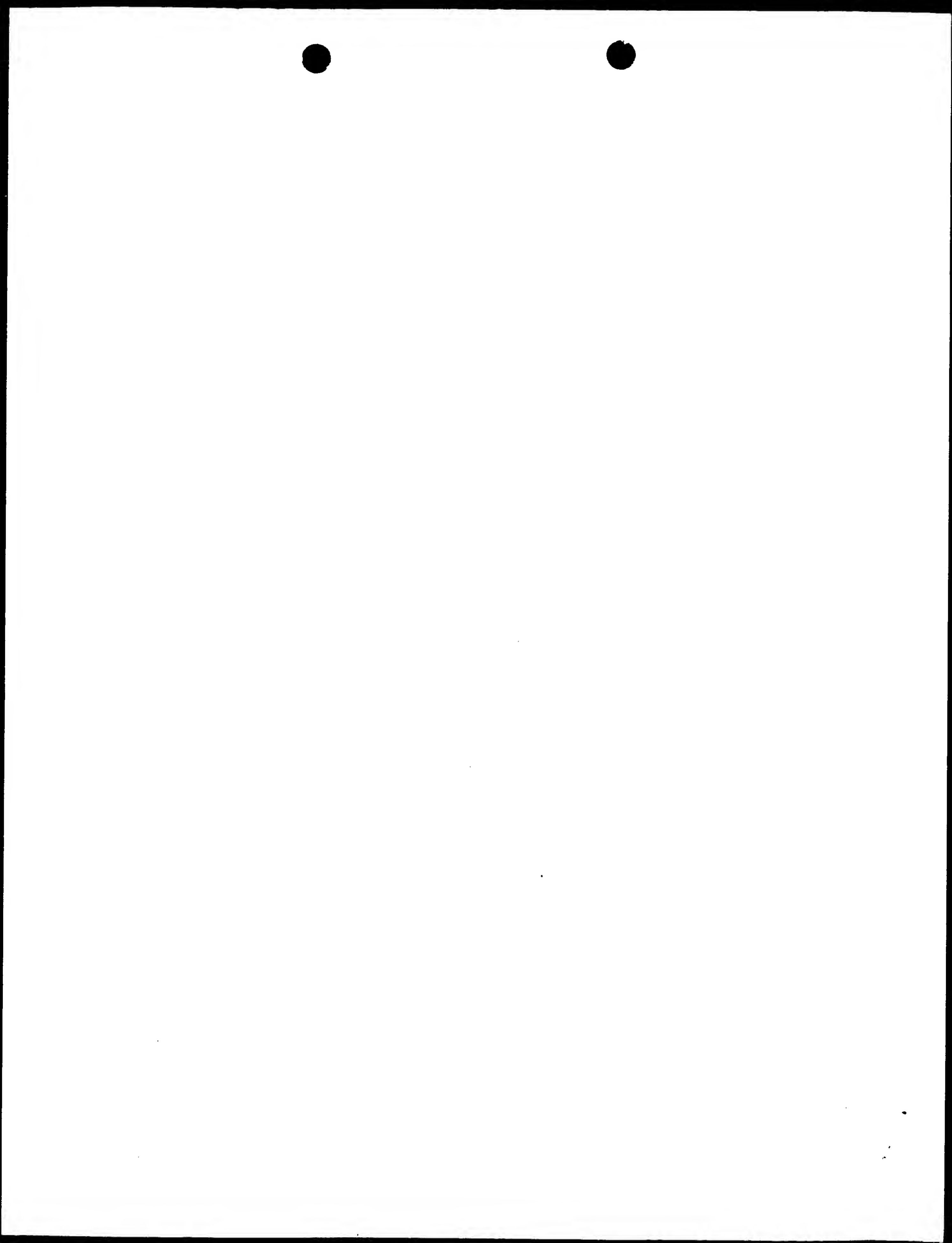
Db 386 AAGAGCCAGTCAAGGTCAGTCTCCACTAAGCCTGGCTCCTGCCCATTTATCTTGATCC 327
QY 94 GTTGGCTATGTTAAACCCACCTAACCGTTGTTTGAAGGACACTGATTGTCCAGGTATCA 153
Db 326 GGTGCGCCATGTTGAATCCCCCTAACCGCTGCTTGAAGATATACTGACTGCCAGGAATCA 267
QY 154 AAAAGTCTGTGAAGGTTCTCTGCGGTATGGCTTGTTCGTTCCACAATAATAG 206
Db 266 AGAAGTCTGTGAAGGCTCTTTCGGGGATGGCTGTTTTCGTTCCCAAGTGAGAG 214

RESULT 75
BF736734/c
LOCUS BF736734 419 bp mRNA linear EST 10-JAN-2001
DEFINITION PM1-KT0004-241100-001-c04 KT0004 Homo sapiens cDNA, mRNA sequence.
ACCESSION BF736734
VERSION BF736734.1 GI:12063408
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 419)
AUTHORS Dias Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Briones,M.R.,
Nagai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.F.,
Goldman,G.H., Carvalho,A.F., Matsukuma,A., Baia,G.S., Simpson,D.H.,
Brunstein,A., deOliveira,P.S., Bucher,P., Jongeneel,C.V., O'Hare
,M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and
Simpson,A.J.
TITLE Shotgun sequencing of the human transcriptome with ORF expressed
sequence tags
JOURNAL Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
MEDLINE 20202663
COMMENT Contact: Simpson A.J.G.
Laboratory of Cancer Genetics
Ludwig Institute for Cancer Research
Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
Brazil
Tel: +55-11-2704922
Fax: +55-11-2707001
Email: asimpson@ludwig.org.br
This sequence was derived from the FAPESP/LICR Human Cancer Genome
Project. This entry can be seen in the following URL
(http://www.ludwig.org.br/scripts/gethtml2.pl?t1=PM1&t2=PM1-KT0004-
241100-001-c04&t3=2000-11-24&t4=1)
Seq primer: puc 18 forward
High quality sequence stop: 419.

FEATURES
source Location/Qualifiers
1..419
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_lib="KT0004"
/dev_stage="Adult"
/note="Organ: bladder tumor; Vector: puc18; Site_1: Smal;
Site_2: Smal; A mini-library was made by cloning products
derived from ORESTES PCR (U.S. Letters Patent application
No. 196,716 - Ludwig Institute for Cancer Research)
profiles into the pUC 18 vector. Reverse transcription of
tissue mRNA and cDNA amplification were performed under
low stringency conditions."
BASE COUNT 99 a 103 c 126 g 91 t
ORIGIN

Query Match 58.9%; Score 121.4; DB 12; Length 419;
Best Local Similarity 79.9%; Pred. No. 3.2e-28;
Matches 143; Conservative 0; Mismatches 36; Indels 0; Gaps 0;
QY 28 ATGCTCAAGAACCCAGTTAAAGGTCCTGTCTACTAAGCCAGGTTCTTGTCTCTATTATCT 87
Db 287 AAGCGCAAGAGCCAGTCAAGGTCCTGCTCCACTAAGCTGGCTCTCTGCCCATTTACCT 228
QY 88 TGATTCTGTCGCTATGTTAAACCCACCTAACCGTTGTTGAAGGACACTGATTGTCCAG 147

Db 227 TGATCCGGTGCGCCGTGTTGAATCCCCCTAACCGCTGCTTGAAGATATACTGACTGCCAG 168
QY 148 GTATCAAAAAGTGTGTGAAGGTTCCCTGCGGTATGGCTTGTTCGTTCCACAATAATAG 206
Db 167 GAATCAAGAAGTGTGTGAAGGCTCTTTCGGGGATGGCTGTTTCGTTCCCAAGTGAGAG 109
Search completed: February 15, 2003, 23:10:13
Job time : 1193 secs



GenCore version 5.1.4 p5_4578
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 15, 2003, 23:10:43 ; Search time 67 Seconds
(without alignments)
1565.980 Million cell updates/sec

Title: 09-833799-13A
Perfect score: 206
Sequence: 1 AATTCGAGCTCGTACCATA.....GTTTCGTTCCACAATAATAG 206

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 424239 seqs, 254661826 residues

Total number of hits satisfying chosen parameters: 848478

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 100 summaries

Database : Published Applications_NA:*

1: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq:*

2: /cgn2_6/ptodata/1/pubpna/PCT_NEW_PUB.seq:*

3: /cgn2_6/ptodata/1/pubpna/US05_NEW_PUB.seq:*

4: /cgn2_6/ptodata/1/pubpna/US06_PUBCOMB.seq:*

5: /cgn2_6/ptodata/1/pubpna/US07_NEW_PUB.seq:*

6: /cgn2_6/ptodata/1/pubpna/PCRTUS_PUBCOMB.seq:*

7: /cgn2_6/ptodata/1/pubpna/US08_NEW_PUB.seq:*

8: /cgn2_6/ptodata/1/pubpna/US08_PUBCOMB.seq:*

9: /cgn2_6/ptodata/1/pubpna/US09_NEW_PUB.seq:*

10: /cgn2_6/ptodata/1/pubpna/US09_PUBCOMB.seq:*

11: /cgn2_6/ptodata/1/pubpna/US10_NEW_PUB.seq:*

12: /cgn2_6/ptodata/1/pubpna/US10_PUBCOMB.seq:*

13: /cgn2_6/ptodata/1/pubpna/US60_NEW_PUB.seq:*

14: /cgn2_6/ptodata/1/pubpna/US60_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
C 1	125.2	60.8	480	10	US-09-920-300A-1393
C 2	125.2	60.8	480	12	US-10-033-528-1393
C 3	123.2	59.8	2309	10	US-09-954-456-1126
C 4	123.2	59.8	2309	10	US-09-954-456-1798
C 5	32.2	15.6	546	10	US-09-815-242-4199
C 6	32.2	15.6	549	10	US-09-815-242-8354
C 7	30.2	14.7	1489	10	US-09-731-872-213
C 8	30.2	14.7	36159	12	US-10-135-687-3
C 9	29.4	14.3	5150	10	US-09-800-528-14
C 10	29	14.1	549	10	US-09-815-242-8688
C 11	29	14.1	549	10	US-09-815-242-8956
C 12	29	14.1	549	10	US-09-925-637-9
C 13	28.6	13.9	180216	10	US-09-835-232-6
C 14	28.4	13.8	464	10	US-09-864-761-2155
C 15	28.4	13.8	1656	10	US-09-729-674-19
C 16	28.4	13.8	1668	10	US-09-765-205-25
C 17	28.4	13.8	99916	10	US-09-816-095-3
C 18	28	13.6	2125	12	US-10-044-090-335
C 19	27.4	13.3	980	10	US-09-785-738A-3

20	27	13.1	354	10	US-09-852-659A-121	Sequence 121, App
21	27	13.1	753	9	US-09-852-797-34	Sequence 34, Appl
22	27	13.1	753	10	US-09-853-161-34	Sequence 34, Appl
23	27	13.1	753	10	US-09-852-659A-34	Sequence 34, Appl
24	27	13.1	783	9	US-09-852-797-50	Sequence 50, Appl
25	27	13.1	783	10	US-09-853-161-50	Sequence 50, Appl
26	27	13.1	783	10	US-09-852-659A-50	Sequence 50, Appl
C 27	27	13.1	1536	10	US-09-815-242-6640	Sequence 6640, Ap
28	26.8	13.0	3976	9	US-10-227-884-59	Sequence 59, Appl
29	26.8	13.0	4531	9	US-10-114-893-118	Sequence 118, App
C 30	26.8	13.0	5686	10	US-09-775-938A-31	Sequence 31, Appl
C 31	26.8	13.0	8210	10	US-09-962-436-279	Sequence 279, App
C 32	26.6	12.9	687	10	US-09-070-927A-612	Sequence 612, App
C 33	26.6	12.9	891	10	US-09-815-242-7107	Sequence 7107, Ap
34	26.6	12.9	1857	10	US-09-974-300-6054	Sequence 6054, Ap
35	26.6	12.9	24004	10	US-09-070-927A-72	Sequence 72, Appl
C 36	26.4	12.8	8921	10	US-09-070-927A-68	Sequence 68, Appl
C 37	26.2	12.7	670	10	US-09-764-869-355	Sequence 355, App
C 38	26.2	12.7	1784	9	US-10-105-930-1	Sequence 1, Appli
C 39	26.2	12.7	2000	9	US-09-938-842A-4875	Sequence 4875, Ap
C 40	26.2	12.7	3479	10	US-09-822-849A-298	Sequence 298, App
C 41	26.2	12.7	157875	9	US-09-935-464-1	Sequence 1, Appli
C 42	26.2	12.7	640681	10	US-09-790-988-1	Sequence 1, Appli
43	26	12.6	1340	9	US-10-214-188-3	Sequence 3, Appli
C 44	26	12.6	2000	9	US-09-938-842A-4030	Sequence 4030, Ap
45	25.8	12.5	340	10	US-09-764-877-2223	Sequence 2223, Ap
46	25.8	12.5	539	10	US-09-070-927A-722	Sequence 722, App
47	25.8	12.5	1047	10	US-09-886-055-410	Sequence 410, App
C 48	25.8	12.5	1049	10	US-09-800-729-67	Sequence 67, Appl
C 49	25.8	12.5	1107	9	US-09-738-626-1385	Sequence 1385, Ap
50	25.8	12.5	1125	9	US-09-965-528-31	Sequence 31, Appl
C 51	25.8	12.5	1600	10	US-09-938-540-1	Sequence 1, Appli
C 52	25.8	12.5	3098	10	US-09-925-300-693	Sequence 693, App
C 53	25.8	12.5	3189	12	US-10-044-090-333	Sequence 333, App
C 54	25.8	12.5	3299	10	US-09-800-729-68	Sequence 68, Appl
C 55	25.8	12.5	4116	9	US-09-822-846-104	Sequence 104, App
C 56	25.8	12.5	6065	10	US-09-800-729-35	Sequence 35, Appl
57	25.6	12.4	247	9	US-10-002-048A-4	Sequence 4, Appli
58	25.6	12.4	327	10	US-09-960-352-6148	Sequence 6148, Ap
C 59	25.6	12.4	434	10	US-09-954-456-2208	Sequence 2208, Ap
C 60	25.6	12.4	434	10	US-09-880-107-2920	Sequence 2920, Ap
61	25.6	12.4	468	10	US-09-783-590-11050	Sequence 11050, A
C 62	25.6	12.4	633	9	US-09-995-898A-29	Sequence 29, Appl
C 63	25.6	12.4	729	10	US-09-910-943-715	Sequence 715, App
64	25.6	12.4	942	10	US-09-815-242-8813	Sequence 8813, Ap
65	25.6	12.4	942	10	US-09-815-242-8969	Sequence 8969, Ap
66	25.6	12.4	1473	9	US-09-995-898A-3	Sequence 3, Appli
67	25.6	12.4	1560	10	US-09-995-898A-28	Sequence 28, Appl
68	25.6	12.4	2211	10	US-09-764-864-359	Sequence 359, App
69	25.6	12.4	6134	10	US-09-764-877-3301	Sequence 3301, Ap
70	25.6	12.4	28690	10	US-09-070-927A-138	Sequence 138, App
C 71	25.4	12.3	326	10	US-09-998-598-2252	Sequence 2252, Ap
C 72	25.4	12.3	327	9	US-09-796-692-7381	Sequence 7381, Ap
C 73	25.4	12.3	504	10	US-09-783-590-7280	Sequence 7280, Ap
74	25.4	12.3	1026	10	US-09-822-830A-205	Sequence 205, App
75	25.4	12.3	1041	8	US-08-927-939-55	Sequence 55, Appl
76	25.4	12.3	1057	9	US-09-974-298-14	Sequence 14, Appl
77	25.4	12.3	1057	12	US-10-044-090-574	Sequence 574, App
78	25.4	12.3	1103	12	US-10-121-119-6	Sequence 6, Appli
79	25.4	12.3	1104	12	US-10-044-090-575	Sequence 575, App
C 80	25.4	12.3	2084	9	US-10-174-590-175	Sequence 175, App
C 81	25.4	12.3	2084	9	US-10-176-758-175	Sequence 175, App
C 82	25.4	12.3	2084	9	US-10-175-737-175	Sequence 175, App
C 83	25.4	12.3	2084	9	US-10-173-706-175	Sequence 175, App
C 84	25.4	12.3	2084	9	US-10-175-738-175	Sequence 175, App
C 85	25.4	12.3	2084	9	US-10-175-752-175	Sequence 175, App
C 86	25.4	12.3	2084	9	US-10-176-482-175	Sequence 175, App
C 87	25.4	12.3	2084	9	US-10-176-757-175	Sequence 175, App
C 88	25.4	12.3	2084	9	US-10-176-913-175	Sequence 175, App
C 89	25.4	12.3	2084	9	US-10-180-552-175	Sequence 175, App
C 90	25.4	12.3	2084	9	US-10-180-557-175	Sequence 175, App
C 91	25.4	12.3	2084	9	US-10-173-700-175	Sequence 175, App
C 92	25.4	12.3	2084	9	US-10-174-572-175	Sequence 175, App

C 93 25.4 12.3 2084 9 US-10-174-579-175 Sequence 175, App
C 94 25.4 12.3 2084 9 US-10-174-582-175 Sequence 175, App
C 95 25.4 12.3 2084 9 US-10-174-588-175 Sequence 175, App
C 96 25.4 12.3 2084 9 US-10-175-739-175 Sequence 175, App
C 97 25.4 12.3 2084 9 US-10-175-740-175 Sequence 175, App
C 98 25.4 12.3 2084 9 US-10-175-743-175 Sequence 175, App
C 99 25.4 12.3 2084 9 US-10-176-488-175 Sequence 175, App
C 100 25.4 12.3 2084 9 US-10-176-492-175 Sequence 175, App

ALIGNMENTS

RESULT 1
US-09-920-300A-1393/c
; Sequence 1393, Application US/09920300A
; Patent No. US20020136728A1
; GENERAL INFORMATION:
; APPLICANT: King, Gordon E.
; APPLICANT: Meagher, Madeleine Joy
; APPLICANT: Xu, Jiangchun
; APPLICANT: Secrist, Heather
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; TITLE OF INVENTION: AND DIAGNOSIS OF COLON CANCER
; FILE REFERENCE: 210121.547
; CURRENT APPLICATION NUMBER: US/09/920,300A
; CURRENT FILING DATE: 2001-07-31
; NUMBER OF SEQ ID NOS: 1789
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 1393
; LENGTH: 480
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: 27, 56, 343
; OTHER INFORMATION: n = A,T,C or G
US-09-920-300A-1393

Query Match 60.8%; Score 125.2; DB 10; Length 480;
Best Local Similarity 81.0%; Pred. No. 2.8e-34;
Matches 145; Conservative 0; Mismatches 34; Indels 0; Gaps 0;

QY 28 ATGCTCAAGAACCCAGTTAAAGGTCCTGTGCTCTACTAAGCCAGGTTCTGTCTCCTATTATCT 87
Db 371 AAGCGCAAGAGCCAGTCAAAGGTCAGTNTCCACTAAGCCTGGCTGCTGCCCATTTATCT 312
QY 88 TGATTGCTTGGCTATGTTAAACCCACCTAACCGTTGTTTGAAGGACACTGATTGTCCAG 147
Db 311 TGATCCGGTGGCCCATGTTGAATCCCCCTAACCGCTGCTTGAAGATACTGACTGCCCCAG 252
QY 148 GTATCAAAAAGTGTGTAAGGTTCTCTGCGGTATGGCTTGTTCGTTCCACAATAATAG 206
Db 251 GAATCAAGAAGTGTGTGAAGGCTCTTGGGGATGGCCTGTTTCGTTCCCCAGTGAGAG 193

RESULT 2
US-10-033-528-1393/c
; Sequence 1393, Application US/10033528
; Patent No. US20020131971A1
; GENERAL INFORMATION:
; APPLICANT: King, Gordon E.
; APPLICANT: Meagher, Madeleine Joy
; APPLICANT: Xu, Jiangchun
; APPLICANT: Secrist, Heather
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; TITLE OF INVENTION: AND DIAGNOSIS OF COLON CANCER
; FILE REFERENCE: 210121.547C1
; CURRENT APPLICATION NUMBER: US/10/033,528
; CURRENT FILING DATE: 2001-12-26
; NUMBER OF SEQ ID NOS: 1896
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 1393

; LENGTH: 480
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: 27, 56, 343
; OTHER INFORMATION: n = A,T,C or G
US-10-033-528-1393

Query Match 60.8%; Score 125.2; DB 12; Length 480;
Best Local Similarity 81.0%; Pred. No. 2.8e-34;
Matches 145; Conservative 0; Mismatches 34; Indels 0; Gaps 0;

QY 28 ATGCTCAAGAACCCAGTTAAAGGTCCTGTGCTCTACTAAGCCAGGTTCTGTCTCCTATTATCT 87
Db 371 AAGCGCAAGAGCCAGTCAAAGGTCAGTNTCCACTAAGCCTGGCTGCTGCCCATTTATCT 312
QY 88 TGATTGCTTGGCTATGTTAAACCCACCTAACCGTTGTTTGAAGGACACTGATTGTCCAG 147
Db 311 TGATCCGGTGGCCCATGTTGAATCCCCCTAACCGCTGCTTGAAGATACTGACTGCCCCAG 252
QY 148 GTATCAAAAAGTGTGTAAGGTTCTCTGCGGTATGGCTTGTTCGTTCCACAATAATAG 206
Db 251 GAATCAAGAAGTGTGTGAAGGCTCTTGGGGATGGCCTGTTTCGTTCCCCAGTGAGAG 193

RESULT 3
US-09-954-456-1126
; Sequence 1126, Application US/09954456
; Patent No. US20020115057A1
; GENERAL INFORMATION:
; APPLICANT: Young, Paul
; TITLE OF INVENTION: Process for Identifying Anti-Cancer Therapeutic Agents Using Can
; TITLE OF INVENTION: Sets
; FILE REFERENCE: 689290-76
; CURRENT APPLICATION NUMBER: US/09/954,456
; CURRENT FILING DATE: 2001-09-18
; PRIOR APPLICATION NUMBER: US/60/233,617
; PRIOR FILING DATE: 2000-09-18
; PRIOR APPLICATION NUMBER: US/60/234,052
; PRIOR FILING DATE: 2000-09-20
; PRIOR APPLICATION NUMBER: US/60/234,923
; PRIOR FILING DATE: 2000-09-25
; PRIOR APPLICATION NUMBER: US/60/235,134
; PRIOR FILING DATE: 2000-09-25
; PRIOR APPLICATION NUMBER: US/60/235,637
; PRIOR FILING DATE: 2000-09-26
; PRIOR APPLICATION NUMBER: US/60/235,638
; PRIOR FILING DATE: 2000-09-26
; PRIOR APPLICATION NUMBER: US/60/235,711
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: US/60/235,720
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: US/60/235,840
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: US/60/235,863
; NUMBER OF SEQ ID NOS: 2276
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 1126
; LENGTH: 2309
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; OTHER INFORMATION: n=a,t,g or c
US-09-954-456-1126

Query Match 59.8%; Score 123.2; DB 10; Length 2309;
Best Local Similarity 81.2%; Pred. No. 2.9e-33;
Matches 143; Conservative 0; Mismatches 33; Indels 0; Gaps 0;

QY 28 ATGCTCAAGAACCCAGTTAAAGGTCCTGTGCTCTACTAAGCCAGGTTCTGTCTCCTATTATCT 87

Db 1552 AAGCGCAAGAGCCAGTCAAGGTCAGTCTCCACTAAGCCTGGCTCCTGCCCATTAATCT 1611
QY 88 TGATTCGTTGCGCTATGTTAAACCCACCTAACCGTTGTTTGAAGGACACTGATTGTCCAG 147
Db 1612 TGATCCGGTGCCTGCTGTTGAATCCCTTAACCGTCTGTTGAAAGATAGTACTGACTGCCAG 1671
QY 148 GTATCAAAAAGTGTGTGAAGGTTCTCTGCGGTATGCTTGTTCGTTCCACAATAA 203
Db 1672 GAATCAAGAAGTGTGTGAAGGCTCTTTCGCGGATGGCTGTTTCGTTCCCAAGTGA 1727

RESULT 4
US-09-954-456-1798
; Sequence 1798, Application US/09954456
; Patent No. US20020115057A1
; GENERAL INFORMATION:
; APPLICANT: Young, Paul
; TITLE OF INVENTION: Process for Identifying Anti-Cancer Therapeutic Agents Using Cand
; FILE OF INVENTION: Sets
; FILE REFERENCE: 689290-76
; CURRENT APPLICATION NUMBER: US/09/954,456
; CURRENT FILING DATE: 2001-09-18
; PRIOR APPLICATION NUMBER: US/60/233,617
; PRIOR FILING DATE: 2000-09-18
; PRIOR APPLICATION NUMBER: US/60/234,052
; PRIOR FILING DATE: 2000-09-20
; PRIOR APPLICATION NUMBER: US/60/234,923
; PRIOR FILING DATE: 2000-09-25
; PRIOR APPLICATION NUMBER: US/60/235,134
; PRIOR FILING DATE: 2000-09-25
; PRIOR APPLICATION NUMBER: US/60/235,637
; PRIOR FILING DATE: 2000-09-26
; PRIOR APPLICATION NUMBER: US/60/235,638
; PRIOR FILING DATE: 2000-09-26
; PRIOR APPLICATION NUMBER: US/60/235,711
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: US/60/235,720
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: US/60/235,840
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: US/60/235,863
; PRIOR FILING DATE: 2000-09-27
; NUMBER OF SEQ ID NOS: 2276
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 1798
; LENGTH: 2309
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; OTHER INFORMATION: n=a,t,g or c
US-09-954-456-1798

Query Match 59.8%; Score 123.2; DB 10; Length 2309;
Best Local Similarity 81.2%; Pred. No. 2.9e-33;
Matches 143; Conservative 0; Mismatches 33; Indels 0; Gaps 0;

QY 28 ATGCTCAAGAACCCAGTTAAAGGTCCTGTGTCTACTAAGCCAGGTTCTGTCTCTATTATCT 87
Db 1552 AAGCGCAAGAGCCAGTCAAGGTCAGTCTCCACTAAGCCTGGCTCCTGCCCATTAATCT 1611
QY 88 TGATTCGTTGCGCTATGTTAAACCCACCTAACCGTTGTTTGAAGGACACTGATTGTCCAG 147
Db 1612 TGATCCGGTGCCTGCTGTTGAATCCCTTAACCGTCTGTTGAAAGATAGTACTGACTGCCAG 1671
QY 148 GTATCAAAAAGTGTGTGAAGGTTCTCTGCGGTATGCTTGTTCGTTCCACAATAA 203
Db 1672 GAATCAAGAAGTGTGTGAAGGCTCTTTCGCGGATGGCTGTTTCGTTCCCAAGTGA 1727

RESULT 5
US-09-815-242-4199/c

; Sequence 4199, Application US/09815242
; Patent No. US20020061569A1
; GENERAL INFORMATION:
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari L.
; APPLICANT: Zyskind, Judith W.
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John D.
; APPLICANT: Carr, Grant J.
; APPLICANT: Yamamoto, Robert T.
; APPLICANT: Xu, H. Howard
; TITLE OF INVENTION: Identification of Essential Genes in
; FILE OF INVENTION: Prokaryotes
; FILE REFERENCE: ELITRA.011A
; CURRENT APPLICATION NUMBER: US/09/815,242
; CURRENT FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; NUMBER OF SEQ ID NOS: 14110
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 4199
; LENGTH: 546
; TYPE: DNA
; ORGANISM: Staphylococcus aureus
US-09-815-242-4199

Query Match 15.6%; Score 32.2; DB 10; Length 546;
Best Local Similarity 48.1%; Pred. No. 0.12;
Matches 91; Conservative 0; Mismatches 98; Indels 0; Gaps 0;

QY 2 ATTCGAGCTCGGTACCATACCTGCATATGCTCAAGAACCCAGTTAAAGTCTGTGTCTAC 61
Db 312 ATTTGGCTTAGACCTGCACCTGCAGAACCTACAAAACCCAGTAACGCTGGTATTCT 253
QY 62 TAAGCCAGGTTCTTGTCTATTATCTTGTTCGCTGCTATGTTAAACCCACCTAACCG 121
Db 252 TACCACATACCATGATTCATCTGTCATGATTATTCCTAAACATATCCAGGAATGT 193
QY 122 TTGTTTGAAGGACACTGATTGTCCAGGTATCAAAAAGTGTGTGAAGTTCTTCCGGTAT 181
Db 192 TTTTAAACGGTTGTTTAGCTTTACCATCTTTAACTTGAGTTTCTTCTTCCGGTAT 133
QY 182 GGCTTGTTT 190
Db 132 GACTACTCT 124

RESULT 6
US-09-815-242-8354/c
; Sequence 8354, Application US/09815242
; Patent No. US20020061569A1
; GENERAL INFORMATION:
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari L.
; APPLICANT: Zyskind, Judith W.
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John D.
; APPLICANT: Carr, Grant J.
; APPLICANT: Yamamoto, Robert T.
; APPLICANT: Xu, H. Howard
; TITLE OF INVENTION: Identification of Essential Genes in

Sun Feb 16 09:12:56 2003

```

; NAME/KEY: CDS
; LOCATION: 460..1290
US-09-731-872-213

Query Match      14.7%; Score 30.2; DB 10; Length 1489;
Best Local Similarity 50.3%; Pred. No. 0.98;
Matches 74; Conservative 0; Mismatches 73; Indels 0; Gaps 0;

QY 35 AGAACCCAGTTAAAGGCTCTGTCTACTAAAGCCAGGTTCTTGTCTCTATTATCTTGTGATTTCG 94
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 145 AGAAAGGGTTAAAGGCTGATGGTACCTAAAGCCCTGGTACTTGAATTTGATCAAGATAAG 204
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

QY 95 TTGGGCTATGTTAAACCCACCTAACCGTTGTTTGAAGGACACTGATTGTCCAGGTATCAA 154
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 205 CTGCCTTAAGTTCTCTTCATTACACAAATGATCCTAGATAATTGATAGATCCTGTGGTTC 264
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

QY 155 AAAGTGCTGTGAAGGTTCTCTGCGGTAT 181
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 265 AACTGGATTCTAGATAGAGAAGCTGGAT 291
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

RESULT 8
US-10-135-687-3/c
; Sequence 3, Application US/10135687
; Patent No. US20020123120A1
; GENERAL INFORMATION:
; APPLICANT: CHANDRAMOULISWARAN, Ishwar et al.
; TITLE OF INVENTION: ISOLATED HUMAN KINASE PROTEINS, NUCLEIC
; TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN KINASE PROTEINS, AND USES
; TITLE OF INVENTION: THEREOF
; FILE REFERENCE: CL001068DIV
; CURRENT APPLICATION NUMBER: US/10/135,687
; CURRENT FILING DATE: 2002-05-01
; PRIOR APPLICATION NUMBER: 09/749,588
; PRIOR FILING DATE: 2000-12-28
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 36159
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(36159)
; OTHER INFORMATION: n = A,T,C or G
US-10-135-687-3

Query Match      14.7%; Score 30.2; DB 12; Length 36159;
Best Local Similarity 49.1%; Pred. No. 4.5;
Matches 80; Conservative 0; Mismatches 83; Indels 0; Gaps 0;

QY 31 CTCAAGAACCAGTTAAAGGTCCTGTGCTACTAAGCCAGGTTCTTGTCTCTATTATCTTGA 90
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 12268 CTCAGATCCAGATCCAGATCCAGATCCAGATCCAGATCCAGATCCAGATCCAGATCCAGAT 12209
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

QY 91 TTCGTTGGGCTATGTTAAACCCACCTAACCGTTGTTTGAAGGACACTGATTGTCCAGGTA 150
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 12208 ACAGATCTTTCCGTGAAGAACCATCTAACTGCGCGTGTCTCTCAGATAATACCCAGGTT 12149
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

QY 151 TCAAAAAGTGCTGCAAGGTTCTCTGCGGTATGGCTTGTTCGT 193
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 12148 TCTCAAACTACTTTCAAGTGATATGTGCTTATTGTTTGT 12106
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

RESULT 9
US-09-800-528-14/c
; Sequence 14, Application US/09800528
; Patent No. US20010016955A1
; GENERAL INFORMATION:
; APPLICANT: Woodhead, Mary Rose
; APPLICANT: Taylor, Mark Andrew
; APPLICANT: Brennan, Rex Michael
; TITLE OF INVENTION: BLACKCURRANT PROMOTERS AND GENES

```

```

; TITLE OF INVENTION: Prokaryotes
; FILE REFERENCE: ELITRA.011A
; CURRENT APPLICATION NUMBER: US/09/815,242
; CURRENT FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; NUMBER OF SEQ ID NOS: 14110
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 8354
; LENGTH: 549
; TYPE: DNA
; ORGANISM: Staphylococcus aureus
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)...(549)
US-09-815-242-8354

Query Match      15.6%; Score 32.2; DB 10; Length 549;
Best Local Similarity 48.1%; Pred. No. 0.12;
Matches 91; Conservative 0; Mismatches 98; Indels 0; Gaps 0;

QY 2 ATTCGAGCTCGGTACCATATGCTCAAGAACCCAGTTAAAGGTCCTGTGCTAC 61
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 312 ATTTGGCTTAGACCTGCACCTGCAGAACCTACAAACCCAGTAACGCGCTGTGATTCTT 253
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

QY 62 TAAGCCAGGTTCTTGTCTCTATTATCTTGTGCTGCTATGTTAAACCCACCTAACCG 121
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 252 TACCACATACCATGATTCTGTCATGATTAAATCCCACTAAACATATCCAGGGAATGT 193
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

QY 122 TTGTTTGAAGGACACTGATTGTCCAGGTATCAAAAAGTGTGTGAAGGTTCTCTGCGGTAT 181
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 192 TTTTAAACGGTTGTTTGTAGCTTTTACCATCTTTAACTTGAGTTTCTTCTTCTCGGTAT 133
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

QY 182 GGCTTGT 190
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 132 GACTACTCT 124
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

RESULT 7
US-09-731-872-213
; Sequence 213, Application US/09731872
; Patent No. US20020102604A1
; GENERAL INFORMATION:
; APPLICANT: Dumas Milne Edwards, Jean Baptiste
; APPLICANT: Bougueleret, Lydie
; APPLICANT: Jobert, Severin
; TITLE OF INVENTION: FULL-LENGTH HUMAN CDNAS ENCODING POTENTIALLY SECRETED PROTEINS
; FILE REFERENCE: 78.US3.REG
; CURRENT APPLICATION NUMBER: US/09/731,872
; CURRENT FILING DATE: 2000-12-07
; PRIOR APPLICATION NUMBER: US 60/169,629
; PRIOR FILING DATE: 1999-12-08
; PRIOR APPLICATION NUMBER: US 60/187,470
; PRIOR FILING DATE: 2000-03-06
; NUMBER OF SEQ ID NOS: 482
; SOFTWARE: Patent.pm
; SEQ ID NO 213
; LENGTH: 1489
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:

```


OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 1.3
OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 1.2
OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 1.3
OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 3.5
OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 1.4
OTHER INFORMATION: EXPRESSED IN BT474, SIGNAL = 1.3
OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 1.6
OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 1.8
OTHER INFORMATION: EXPRESSED IN HBL100, SIGNAL = 1.5
OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 1.4
US-09-864-761-2155

Query Match 13.8%; Score 28.4; DB 10; Length 464;
Best Local Similarity 54.9%; Pred. No. 2.4;
Matches 56; Conservative 0; Mismatches 46; Indels 0; Gaps 0;

QY 77 TCCTATTATCTTGATTGCTGCGCTATGTTAAACCCACCTAACCGTTGTTGAAGGACAC 136
|||||
Db 293 TCCTCTTCCACAGATTATTGCTGCTCCTCCACAGCCCTTCACTGTAGATGGAAGCCTG 352
|||||
QY 137 TGATTGTCCAGGTATCAAAAAGTGTGTGAAGTTCCTGCGG 178
|||||
Db 353 TGATTTCCTGGCATCATGCATGACTGAGAAAGGTCAGTGTGG 394
|||||

RESULT 15

US-09-729-674-19/c
Sequence 19, Application US/09729674
Patent No. US20010039335A1
GENERAL INFORMATION:
APPLICANT: Jacobs, Kenneth
APPLICANT: McCoy, John M.
APPLICANT: LaVallie, Edward R.
APPLICANT: Collins-Racie, Lisa A.
APPLICANT: Evans, Cheryl
APPLICANT: Merberg, David
APPLICANT: Treacy, Maurice
APPLICANT: Agostino, Michael J.
APPLICANT: Steininger II, Robert J.
APPLICANT: Spaulding, Vikki
APPLICANT: Wong, Gordon G.
APPLICANT: Clark, Hilary
APPLICANT: Fechtel, Kim
APPLICANT: Genetics Institute, Inc.
TITLE OF INVENTION: SECRETED PROTEINS AND POLYNUCLEOTIDES ENCODING THEM
FILE REFERENCE: 6055-64X
CURRENT APPLICATION NUMBER: US/09/729,674
CURRENT FILING DATE: 2000-12-04
PRIOR APPLICATION NUMBER: 09/539,330
PRIOR FILING DATE: 2000-03-30
NUMBER OF SEQ ID NOS: 283
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 19
LENGTH: 1656
TYPE: DNA
ORGANISM: Homo sapiens
US-09-729-674-19

Query Match 13.8%; Score 28.4; DB 10; Length 1656;
Best Local Similarity 62.9%; Pred. No. 4.4;
Matches 44; Conservative 0; Mismatches 26; Indels 0; Gaps 0;

QY 32 TCAAGAACCAAGTTAAAGTCTGTGCTACTAAGCCAGGTTCTTGTCTTATTTATCTTGAT 91
|||||
Db 1023 TCAAGCTACAGGTAAGGCTCTACCACCATTAAGGTATTGACTTATCTATGAATTTGAT 964
|||||
QY 92 TCGTTGCGCT 101
|||||
Db 963 TCTTCTTCT 954
|||||

RESULT 16
US-09-765-205-25/c

Sequence 25, Application US/09765205
Patent No. US20020034800A1
GENERAL INFORMATION:
APPLICANT: Cao, Li
TITLE OF INVENTION: BONE MARROW SECRETED PROTEINS AND POLYNUCLEOTIDES
FILE REFERENCE: 1458.004/200130.449
CURRENT APPLICATION NUMBER: US/09/765,205
CURRENT FILING DATE: 2001-01-17
PRIOR APPLICATION NUMBER: US/09/212,440
PRIOR FILING DATE: 1998-12-16
NUMBER OF SEQ ID NOS: 46
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 25
LENGTH: 1668
TYPE: DNA
ORGANISM: human
US-09-765-205-25

Query Match 13.8%; Score 28.4; DB 10; Length 1668;
Best Local Similarity 62.9%; Pred. No. 4.4;
Matches 44; Conservative 0; Mismatches 26; Indels 0; Gaps 0;

QY 32 TCAAGAACCAAGTTAAAGTCTGTGCTACTAAGCCAGGTTCTTGTCTTATTTATCTTGAT 91
|||||
Db 1094 TCAAGCTACAGGTAAGGCTCTACCACCATTAAGGTATTGACTTATCTATGAATTTGAT 1035
|||||
QY 92 TCGTTGCGCT 101
|||||
Db 1034 TCTTCTTCT 1025
|||||

RESULT 17

US-09-816-095-3
Sequence 3, Application US/09816095
Patent No. US20020137164A1
GENERAL INFORMATION:
APPLICANT: GAN, Weiniu
TITLE OF INVENTION: ISOLATED HUMAN ENZYME PROTEINS, NUCLEIC
TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN ENZYME PROTEINS, AND USES
TITLE OF INVENTION: THEREOF
FILE REFERENCE: CL001147
CURRENT APPLICATION NUMBER: US/09/816,095
CURRENT FILING DATE: 2001-03-26
NUMBER OF SEQ ID NOS: 5
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 3
LENGTH: 99916
TYPE: DNA
ORGANISM: Human
FEATURE:
NAME/KEY: misc_feature
LOCATION: (1)...(99916)
OTHER INFORMATION: n = A,T,C or G
US-09-816-095-3

Query Match 13.8%; Score 28.4; DB 10; Length 99916;
Best Local Similarity 51.6%; Pred. No. 31;
Matches 65; Conservative 0; Mismatches 61; Indels 0; Gaps 0;

QY 77 TCCTATTATCTTGATTGCTGCGTATGTTAAACCCACCTAACCGTTGTTGAAGGACAC 136
|||||
Db 75420 TACTATTTTCTACAGCAGCTGCATCAATTTTACACTCTCCACAGCAATGCAGGAAGGTTCC 75479
|||||
QY 137 TGATTGTCCAGGTATCAAAAAGTGTGTGAAGTTCCTGCGGTATGGCTTGTTCGTTCC 196
|||||
Db 75480 AGATTCTCCATATCTCTCACTAATACTATTATTGCTTCTGTTTGTGTTAGTTTATTA 75539
|||||
QY 197 ACAATA 202
|||||
Db 75540 ATAATA 75545
|||||

RESULT 18

US-10-044-090-335/c
; Sequence 335, Application US/10044090
; Patent No. US20020137081A1
; GENERAL INFORMATION:
; APPLICANT: Olga Bandman
; TITLE OF INVENTION: GENES DIFFERENTIALLY EXPRESSED IN VASCULAR TISSUE ACTIVATION
; FILE REFERENCE: PA-0028 US
; CURRENT APPLICATION NUMBER: US/10/044,090
; CURRENT FILING DATE: 2002-01-09
; NUMBER OF SEQ ID NOS: 850
; SOFTWARE: PERL Program
; SEQ ID NO 335
; LENGTH: 2125
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; OTHER INFORMATION: Incyte ID No. US20020137081A1 1989707CB1
US-10-044-090-335

	Query Match	13.6%	Score 28;	DB 12;	Length 2125;
	Best Local Similarity	56.5%;	Pred. No. 6.8;		
	Matches 52; Conservative	0;	Mismatches 40;	Indels 0;	Gaps 0;
Oy	112	CACCTAACCGTGTGTTGAAGGCACACTGATTGTCCAGGTATCAAAAAGTCCTGTGAAGGTT	171		
Dd	1275	CAACTGGTCGTTGAGGAGGCCAGGCAGGTCACAGAGCCAGACATGGCTGGGCCGGTT	1216		
Oy	172	CCTGGCGGTATGGCTTGTTTCGTTCCACAATAA	203		
Dd	1215	CCTGCCCATCTGGGGTCCTCCACTCCACAATCA	1184		

RESULT 19
 US-09-785-738A-3/c
 ; Sequence 3, Application US/09785738A
 ; Patent No. US20020032918A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Sauter, Margret M
 ; APPLICANT: Lorbiecke, Rene
 ; TITLE OF INVENTION: ALTERATION OF GROWTH AND ADAPTATION UNDER HYPOXIC
 ; TITLE OF INVENTION: CONDITIONS
 ; FILE REFERENCE: 2283/201
 ; CURRENT APPLICATION NUMBER: US/09/785,738A
 ; CURRENT FILING DATE: 2001-02-16
 ; PRIOR APPLICATION NUMBER: 60/183,572
 ; PRIOR FILING DATE: 2000-02-18
 ; NUMBER OF SEQ ID NOS: 18
 ; SOFTWARE: PatentIn ver. 2.1
 ; SEQ ID NO 3
 ; LENGTH: 980
 ; TYPE: DNA
 ; ORGANISM: Rice
 ; FEATURE:
 ; NAME/KEY: CDS
 ; LOCATION: (139)..(735)
 ; US-09-785-738A-3

[illegible]

RESULT 20
US-09-852-659A-121
; Sequence 121, Application US/09852659A
; Patent No. US2002007287A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: 28 Human Secreted Proteins
; FILE REFERENCE: PZ003P4
; CURRENT APPLICATION NUMBER: US/09/852,659A
; CURRENT FILING DATE: 2001-05-11
; PRIOR APPLICATION NUMBER: 60/265,583
; PRIOR FILING DATE: 2001-02-02
; PRIOR APPLICATION NUMBER: 09/152,060
; PRIOR FILING DATE: 1998-09-11
; PRIOR APPLICATION NUMBER: PCT/US98/04858
; PRIOR FILING DATE: 1998-03-12
; PRIOR APPLICATION NUMBER: 60/040,762
; PRIOR FILING DATE: 1997-03-14
; PRIOR APPLICATION NUMBER: 60/040,710
; PRIOR FILING DATE: 1997-03-14
; PRIOR APPLICATION NUMBER: 60/050,934
; PRIOR FILING DATE: 1997-05-30
; PRIOR APPLICATION NUMBER: 60/048,100
; PRIOR FILING DATE: 1997-05-30
; PRIOR APPLICATION NUMBER: 60/048,357
; PRIOR FILING DATE: 1997-05-30
; PRIOR APPLICATION NUMBER: 60/048,189
; PRIOR FILING DATE: 1997-05-30
; PRIOR APPLICATION NUMBER: 60/057,765
; PRIOR FILING DATE: 1997-09-05
; PRIOR APPLICATION NUMBER: 60/048,970
; PRIOR FILING DATE: 1997-06-06
; PRIOR APPLICATION NUMBER: 60/068,368
; PRIOR FILING DATE: 1997-12-19
; NUMBER OF SEQ ID NOS: 121
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 121
; LENGTH: 354
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-852-659A-121

[illegible]

```

RESULT 21
US-09-852-797-34
; Sequence 34, Application US/09852797
; Patent No. US20020172994A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: 28 Human Secreted Proteins
; FILE REFERENCE: PZ003P2
; CURRENT APPLICATION NUMBER: US/09/852,797
; CURRENT FILING DATE: 2001-05-11
; PRIOR APPLICATION NUMBER: 60/265,583
; PRIOR FILING DATE: 2001-02-02
; PRIOR APPLICATION NUMBER: 09/152,060
; PRIOR FILING DATE: 1998-09-11
; PRIOR APPLICATION NUMBER: PCT/US98/04858
; PRIOR FILING DATE: 1998-03-12
;

```

; PRIOR APPLICATION NUMBER: 60/040,762
; PRIOR FILING DATE: 1997-03-14
; PRIOR APPLICATION NUMBER: 60/040,710
; PRIOR FILING DATE: 1997-03-14
; PRIOR APPLICATION NUMBER: 60/050,934
; PRIOR FILING DATE: 1997-05-30
; PRIOR APPLICATION NUMBER: 60/048,100
; PRIOR FILING DATE: 1997-05-30
; PRIOR APPLICATION NUMBER: 60/048,357
; PRIOR FILING DATE: 1997-05-30
; PRIOR APPLICATION NUMBER: 60/048,189
; PRIOR FILING DATE: 1997-05-30
; PRIOR APPLICATION NUMBER: 60/057,765
; PRIOR FILING DATE: 1997-09-05
; PRIOR APPLICATION NUMBER: 60/048,970
; PRIOR FILING DATE: 1997-06-06
; PRIOR APPLICATION NUMBER: 60/068,368
; PRIOR FILING DATE: 1997-12-19
; NUMBER OF SEQ ID NOS: 118
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 34
; LENGTH: 753
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-852-797-34

Query Match 13.1%; Score 27; DB 9; Length 753;
Best Local Similarity 57.8%; Pred. No. 9.3;
Matches 48; Conservative 0; Mismatches 35; Indels 0; Gaps 0;

Qy 123 TGTGTTGAAGGACACTGATTGTCCAGGTATCAAAAAGTCTGTGAAGGTTCTCTGCGGTATG 182
Db 172 TGTACAAAGGACAGACAATGCCAGGACAACAAGAAGTGTGTCTTTCAGCTGCGGAAAA 231
Qy 183 GCTTGTTCGTTCCACAATAATA 205
Db 232 AAATGTTTAGATCTCAACAAGA 254

RESULT 22

US-09-853-161-34
; Sequence 34, Application US/09853161
; Patent No. US20020076756A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: 28 Human Secreted Proteins
; FILE REFERENCE: PZ003P3
; CURRENT APPLICATION NUMBER: US/09/853,161
; CURRENT FILING DATE: 2001-05-11
; PRIOR APPLICATION NUMBER: 60/265,583
; PRIOR FILING DATE: 2001-02-02
; PRIOR APPLICATION NUMBER: 60/152,060
; PRIOR FILING DATE: 1998-09-11
; PRIOR APPLICATION NUMBER: PCT/US98/04858
; PRIOR FILING DATE: 1998-03-12
; PRIOR APPLICATION NUMBER: 60/040,762
; PRIOR FILING DATE: 1997-03-14
; PRIOR APPLICATION NUMBER: 60/050,934
; PRIOR FILING DATE: 1997-05-30
; PRIOR APPLICATION NUMBER: 60/048,100
; PRIOR FILING DATE: 1997-05-30
; PRIOR APPLICATION NUMBER: 60/048,357
; PRIOR FILING DATE: 1997-05-30
; PRIOR APPLICATION NUMBER: 60/048,189
; PRIOR FILING DATE: 1997-05-30
; PRIOR APPLICATION NUMBER: 60/057,765
; PRIOR FILING DATE: 1997-09-05
; PRIOR APPLICATION NUMBER: 60/048,970
; PRIOR FILING DATE: 1997-06-06
; PRIOR APPLICATION NUMBER: 60/068,368
; PRIOR FILING DATE: 1997-12-19

; NUMBER OF SEQ ID NOS: 118
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 34
; LENGTH: 753
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-853-161-34

Query Match 13.1%; Score 27; DB 10; Length 753;
Best Local Similarity 57.8%; Pred. No. 9.3;
Matches 48; Conservative 0; Mismatches 35; Indels 0; Gaps 0;

Qy 123 TGTGTTGAAGGACACTGATTGTCCAGGTATCAAAAAGTCTGTGAAGGTTCTCTGCGGTATG 182
Db 172 TGTACAAAGGACAGACAATGCCAGGACAACAAGAAGTGTGTCTTTCAGCTGCGGAAAA 231
Qy 183 GCTTGTTCGTTCCACAATAATA 205
Db 232 AAATGTTTAGATCTCAACAAGA 254

RESULT 23

US-09-852-659A-34
; Sequence 34, Application US/09852659A
; Patent No. US20020077287A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: 28 Human Secreted Proteins
; FILE REFERENCE: PZ003P4
; CURRENT APPLICATION NUMBER: US/09/852,659A
; CURRENT FILING DATE: 2001-05-11
; PRIOR APPLICATION NUMBER: 60/265,583
; PRIOR FILING DATE: 2001-02-02
; PRIOR APPLICATION NUMBER: 60/152,060
; PRIOR FILING DATE: 1998-09-11
; PRIOR APPLICATION NUMBER: PCT/US98/04858
; PRIOR FILING DATE: 1998-03-12
; PRIOR APPLICATION NUMBER: 60/040,762
; PRIOR FILING DATE: 1997-03-14
; PRIOR APPLICATION NUMBER: 60/040,710
; PRIOR FILING DATE: 1997-03-14
; PRIOR APPLICATION NUMBER: 60/050,934
; PRIOR FILING DATE: 1997-05-30
; PRIOR APPLICATION NUMBER: 60/048,100
; PRIOR FILING DATE: 1997-05-30
; PRIOR APPLICATION NUMBER: 60/048,357
; PRIOR FILING DATE: 1997-05-30
; PRIOR APPLICATION NUMBER: 60/048,189
; PRIOR FILING DATE: 1997-05-30
; PRIOR APPLICATION NUMBER: 60/057,765
; PRIOR FILING DATE: 1997-09-05
; PRIOR APPLICATION NUMBER: 60/048,970
; PRIOR FILING DATE: 1997-06-06
; PRIOR APPLICATION NUMBER: 60/068,368
; PRIOR FILING DATE: 1997-12-19
; NUMBER OF SEQ ID NOS: 121
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 34
; LENGTH: 753
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-852-659A-34

Query Match 13.1%; Score 27; DB 10; Length 753;
Best Local Similarity 57.8%; Pred. No. 9.3;
Matches 48; Conservative 0; Mismatches 35; Indels 0; Gaps 0;

Qy 123 TGTGTTGAAGGACACTGATTGTCCAGGTATCAAAAAGTCTGTGAAGGTTCTCTGCGGTATG 182
Db 172 TGTACAAAGGACAGACAATGCCAGGACAACAAGAAGTGTGTCTTTCAGCTGCGGAAAA 231
Qy 183 GCTTGTTCGTTCCACAATAATA 205

Db 232 AAATGTTTAGATCTCAACAAGA 254

RESULT 24

US-09-852-797-50

; Sequence 50, Application US/09852797

; Patent No. US20020172994A1

; GENERAL INFORMATION:

; APPLICANT: Rosen et al.

; TITLE OF INVENTION: 28 Human Secreted Proteins

; FILE REFERENCE: PZ003P2

; CURRENT APPLICATION NUMBER: US/09/852,797

; CURRENT FILING DATE: 2001-05-11

; PRIOR APPLICATION NUMBER: 60/265,583

; PRIOR FILING DATE: 2001-02-02

; PRIOR APPLICATION NUMBER: 09/152,060

; PRIOR FILING DATE: 1998-09-11

; PRIOR APPLICATION NUMBER: PCT/US98/04858

; PRIOR FILING DATE: 1998-03-12

; PRIOR APPLICATION NUMBER: 60/040,762

; PRIOR FILING DATE: 1997-03-14

; PRIOR APPLICATION NUMBER: 60/040,710

; PRIOR FILING DATE: 1997-03-14

; PRIOR APPLICATION NUMBER: 60/050,934

; PRIOR FILING DATE: 1997-05-30

; PRIOR APPLICATION NUMBER: 60/048,100

; PRIOR FILING DATE: 1997-05-30

; PRIOR APPLICATION NUMBER: 60/048,357

; PRIOR FILING DATE: 1997-05-30

; PRIOR APPLICATION NUMBER: 60/057,765

; PRIOR FILING DATE: 1997-09-05

; PRIOR APPLICATION NUMBER: 60/048,970

; PRIOR FILING DATE: 1997-06-06

; PRIOR APPLICATION NUMBER: 60/068,368

; PRIOR FILING DATE: 1997-12-19

; NUMBER OF SEQ ID NOS: 118

; SOFTWARE: PatentIn Ver. 2.0

; SEQ ID NO 50

; LENGTH: 783

; TYPE: DNA

; ORGANISM: Homo sapiens

US-09-852-797-50

Query Match 13.1%; Score 27; DB 9; Length 783;

Best Local Similarity 57.8%; Pred. No. 9.5;

Matches 48; Conservative 0; Mismatches 35; Indels 0; Gaps 0;

QY 123 TGTGTTGAAGGACACTGATTGTCAGGTATCAAAAAGTGTGTGAAGTTCTCTCGGGTATG 182

Db 172 TGTACAAAGGACAGACAATGCCAGGACAAACAAGAAGTGTGTCTTCAGCTCGGAAAA 231

QY 183 GCTTGTTCGTTCCACAATAATA 205

Db 232 AAATGTTTAGATCTCAACAAGA 254

RESULT 25

US-09-853-161-50

; Sequence 50, Application US/09853161

; Patent No. US20020076756A1

; GENERAL INFORMATION:

; APPLICANT: Rosen et al.

; TITLE OF INVENTION: 28 Human Secreted Proteins

; FILE REFERENCE: PZ003P3

; CURRENT APPLICATION NUMBER: US/09/853,161

; CURRENT FILING DATE: 2001-05-11

; PRIOR APPLICATION NUMBER: 60/265,583

; PRIOR FILING DATE: 2001-02-02

; PRIOR APPLICATION NUMBER: 09/152,060

; PRIOR FILING DATE: 1998-09-11

; PRIOR APPLICATION NUMBER: PCT/US98/04858

; PRIOR FILING DATE: 1998-03-12

; PRIOR APPLICATION NUMBER: 60/040,762

; PRIOR FILING DATE: 1997-03-14

; PRIOR APPLICATION NUMBER: 60/040,710

; PRIOR FILING DATE: 1997-03-14

; PRIOR APPLICATION NUMBER: 60/050,934

; PRIOR FILING DATE: 1997-05-30

; PRIOR APPLICATION NUMBER: 60/048,100

; PRIOR FILING DATE: 1997-05-30

; PRIOR APPLICATION NUMBER: 60/048,357

; PRIOR FILING DATE: 1997-05-30

; PRIOR APPLICATION NUMBER: 60/048,189

; PRIOR FILING DATE: 1997-05-30

; PRIOR APPLICATION NUMBER: 60/057,765

; PRIOR FILING DATE: 1997-09-05

; PRIOR APPLICATION NUMBER: 60/048,970

; PRIOR FILING DATE: 1997-06-06

; PRIOR APPLICATION NUMBER: 60/068,368

; PRIOR FILING DATE: 1997-12-19

; NUMBER OF SEQ ID NOS: 118

; SOFTWARE: PatentIn Ver. 2.0

; SEQ ID NO 50

; LENGTH: 783

; TYPE: DNA

; ORGANISM: Homo sapiens

US-09-853-161-50

Query Match 13.1%; Score 27; DB 10; Length 783;

Best Local Similarity 57.8%; Pred. No. 9.5;

Matches 48; Conservative 0; Mismatches 35; Indels 0; Gaps 0;

QY 123 TGTGTTGAAGGACACTGATTGTCAGGTATCAAAAAGTGTGTGAAGTTCTCTCGGGTATG 182

Db 172 TGTACAAAGGACAGACAATGCCAGGACAAACAAGAAGTGTGTCTTCAGCTCGGAAAA 231

QY 183 GCTTGTTCGTTCCACAATAATA 205

Db 232 AAATGTTTAGATCTCAACAAGA 254

RESULT 26

US-09-852-659A-50

; Sequence 50, Application US/09852659A

; Patent No. US20020077287A1

; GENERAL INFORMATION:

; APPLICANT: Rosen et al.

; TITLE OF INVENTION: 28 Human Secreted Proteins

; FILE REFERENCE: PZ003P4

; CURRENT APPLICATION NUMBER: US/09/852,659A

; CURRENT FILING DATE: 2001-05-11

; PRIOR APPLICATION NUMBER: 60/265,583

; PRIOR FILING DATE: 2001-02-02

; PRIOR APPLICATION NUMBER: 09/152,060

; PRIOR FILING DATE: 1998-09-11

; PRIOR APPLICATION NUMBER: PCT/US98/04858

; PRIOR FILING DATE: 1998-03-12

; PRIOR APPLICATION NUMBER: 60/040,762

; PRIOR FILING DATE: 1997-03-14

; PRIOR APPLICATION NUMBER: 60/040,710

; PRIOR FILING DATE: 1997-03-14

; PRIOR APPLICATION NUMBER: 60/050,934

; PRIOR FILING DATE: 1997-05-30

; PRIOR APPLICATION NUMBER: 60/048,100

; PRIOR FILING DATE: 1997-05-30

; PRIOR APPLICATION NUMBER: 60/048,357

; PRIOR FILING DATE: 1997-05-30

; PRIOR APPLICATION NUMBER: 60/048,189

; PRIOR FILING DATE: 1997-05-30

; PRIOR APPLICATION NUMBER: 60/057,765

; PRIOR FILING DATE: 1997-09-05

; PRIOR APPLICATION NUMBER: 60/048,970

; PRIOR FILING DATE: 1997-06-06

; PRIOR APPLICATION NUMBER: 60/068,368

Db 1455 TAAAGTTCCTTTCTAGCAGCTGCTGTTCTTGTGCTATCTTGTGTCATTTCCGAC 1396

Db 1455 TAAAGTTCCTTTCTAGCAGCTGCTGTTCTTGTGTAATCTTGTGTCATTTCCGAC 1396

PRIOR FILING DATE: 1998-06-25	PRIOR APPLICATION NUMBER: 60/091982
PRIOR FILING DATE: 1998-07-07	PRIOR APPLICATION NUMBER: 60/095302
PRIOR FILING DATE: 1998-08-04	PRIOR APPLICATION NUMBER: 60/095318
PRIOR FILING DATE: 1998-08-04	PRIOR APPLICATION NUMBER: 60/095318
PRIOR FILING DATE: 1998-08-10	PRIOR APPLICATION NUMBER: 60/095916
PRIOR FILING DATE: 1998-08-10	PRIOR APPLICATION NUMBER: 60/096146
PRIOR FILING DATE: 1998-08-11	PRIOR APPLICATION NUMBER: 60/096791
PRIOR FILING DATE: 1998-08-17	PRIOR APPLICATION NUMBER: 60/097986
PRIOR FILING DATE: 1998-08-26	PRIOR APPLICATION NUMBER: 60/098544
PRIOR FILING DATE: 1998-08-31	PRIOR APPLICATION NUMBER: 60/099596
PRIOR FILING DATE: 1998-09-09	PRIOR APPLICATION NUMBER: 60/099598
PRIOR FILING DATE: 1998-09-09	PRIOR APPLICATION NUMBER: 60/099803
PRIOR FILING DATE: 1998-09-10	PRIOR APPLICATION NUMBER: 60/099811
PRIOR FILING DATE: 1998-09-10	PRIOR APPLICATION NUMBER: 60/099812
PRIOR FILING DATE: 1998-09-10	PRIOR APPLICATION NUMBER: 60/099816
PRIOR FILING DATE: 1998-09-10	PRIOR APPLICATION NUMBER: 60/100038
PRIOR FILING DATE: 1998-09-11	PRIOR APPLICATION NUMBER: 60/100385
PRIOR FILING DATE: 1998-09-15	PRIOR APPLICATION NUMBER: 60/100390
PRIOR FILING DATE: 1998-09-15	PRIOR APPLICATION NUMBER: 60/100627
PRIOR FILING DATE: 1998-09-16	PRIOR APPLICATION NUMBER: 60/100848
PRIOR FILING DATE: 1998-09-18	PRIOR APPLICATION NUMBER: 60/100919
PRIOR FILING DATE: 1998-09-17	PRIOR APPLICATION NUMBER: 60/101477
PRIOR FILING DATE: 1998-09-23	PRIOR APPLICATION NUMBER: 60/101738
PRIOR FILING DATE: 1998-09-24	PRIOR APPLICATION NUMBER: 60/101741
PRIOR FILING DATE: 1998-09-24	PRIOR APPLICATION NUMBER: 60/101786
PRIOR FILING DATE: 1998-09-25	PRIOR APPLICATION NUMBER: 60/101916
PRIOR FILING DATE: 1998-09-24	PRIOR APPLICATION NUMBER: 60/101922
PRIOR FILING DATE: 1998-09-24	PRIOR APPLICATION NUMBER: 60/106178
PRIOR FILING DATE: 1998-10-28	PRIOR APPLICATION NUMBER: 60/106248
PRIOR FILING DATE: 1998-10-29	PRIOR APPLICATION NUMBER: 60/106464
PRIOR FILING DATE: 1998-10-30	PRIOR APPLICATION NUMBER: 60/106905
PRIOR FILING DATE: 1998-11-03	PRIOR APPLICATION NUMBER: 60/108787
PRIOR FILING DATE: 1998-11-17	PRIOR APPLICATION NUMBER: 60/108801
PRIOR FILING DATE: 1998-11-17	PRIOR APPLICATION NUMBER: 60/108849
PRIOR FILING DATE: 1998-11-18	PRIOR APPLICATION NUMBER: 60/112422
PRIOR FILING DATE: 1998-12-15	PRIOR APPLICATION NUMBER: 60/113296
PRIOR FILING DATE: 1998-12-22	PRIOR APPLICATION NUMBER: 60/113605
PRIOR FILING DATE: 1998-12-23	PRIOR APPLICATION NUMBER: 60/113605

[illegible]

QY 157 AGTGCT 162
|||
Db 3053 AGAGGT 3058

RESULT 29

US-10-114-893-118
; Sequence 118, Application US/10114893
; Publication No. US20020193567A1

GENERAL INFORMATION:

; APPLICANT: Jacobs, Kenneth
; APPLICANT: McCoy, John M.
; APPLICANT: Lavallie, Edward R.
; APPLICANT: Collins-Racie, Lisa A.
; APPLICANT: Evans, Cheryl
; APPLICANT: Merberg, David
; APPLICANT: Treacy, Maurice
; APPLICANT: Bowman, Michael R.
; APPLICANT: Spaulding, Vikki
; APPLICANT: Carlin-Duckett, McKeough
; APPLICANT: Kelleher, Kerry S.
; APPLICANT: Genetics Institute, Inc.
; TITLE OF INVENTION: SECRETED PROTEINS AND POLYNUCLEOTIDES ENCODING THEM
; FILE REFERENCE: GI 6000-10A
; CURRENT APPLICATION NUMBER: US/10/114,893
; CURRENT FILING DATE: 2002-04-02
; EARLIER APPLICATION NUMBER: 09/413,232
; EARLIER FILING DATE: 1999-10-06
; NUMBER OF SEQ ID NOS: 321
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 118
; LENGTH: 4531
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-114-893-118

Query Match 13.0%; Score 26.8; DB 9; Length 4531;
Best Local Similarity 50.8%; Pred. No. 26;
Matches 64; Conservative 0; Mismatches 62; Indels 0; Gaps 0;

QY 37 AACCAAGTTAAAGGTCCTGTCTACTAAGCCAGGTTCTTGCTCCTATTATCTTGATTCGTT 96
|||
Db 3470 AATAATTCAAGTTGAAGTGTGTTTAAACACITTTGTTGTAATGTTTGAATCTCTT 3529
|||
QY 97 GCGCTATGTTAAACCCACCTAACCGTTGTTTGAAGGACACTGATTGTCCAGGTATCAAAA 156
|||
Db 3530 GAGATGTGTTTACCCCACTAGATACATATTTGCCACCTGGTAGTTCTCCATCTAAGCTCA 3589
|||
QY 157 AGTGCT 162
|||
Db 3590 AGAGGT 3595

RESULT 30

US-09-775-938A-31/c
; Sequence 31, Application US/09775938A
; Patent No. US20020081665A1

GENERAL INFORMATION:

; APPLICANT: Haygood, M.
; APPLICANT: Davidson, S.K.
; APPLICANT: Allen, S.W.
; APPLICANT: Hildebrand, M.
; TITLE OF INVENTION: Bryostatins, Bryopyrans and Polyketides: Compositions and Methods
; FILE REFERENCE: 1133.010US1
; CURRENT APPLICATION NUMBER: US/09/775,938A
; CURRENT FILING DATE: 2001-08-30
; PRIOR APPLICATION NUMBER: PCT/US00/21326
; PRIOR FILING DATE: 2000-08-04
; PRIOR APPLICATION NUMBER: US 60/147,283
; PRIOR FILING DATE: 1999-08-04
; NUMBER OF SEQ ID NOS: 38
; SOFTWARE: FastSEQ for Windows Version 4.0

; SEQ ID NO 31
; LENGTH: 5686

; TYPE: DNA
; ORGANISM: Endobugula sertula

FEATURE:

; NAME/KEY: misc_feature
; LOCATION: (1)...(5686)
; OTHER INFORMATION: N refers to any nucleotide.
US-09-775-938A-31

Query Match 13.0%; Score 26.8; DB 10; Length 5686;
Best Local Similarity 50.0%; Pred. No. 29;
Matches 49; Conservative 6; Mismatches 43; Indels 0; Gaps 0;

QY 30 GCTCAAGAACCAAGTTAAAGGTCCTGTGTCTACTAAGCCAGGTTCTTGCTCCTATTATCTTG 89
|||
Db 2495 GCCCAAGTTGCAATGTWTCGGCAARAWCAATTAAGCSMAWTCCCTACCTYTTTGCTTC 2436
|||
QY 90 ATTCGTTGCGCTATGTTAAACCCACCTAACCGTTGTTT 127
|||
Db 2435 ATTTTGTGATAAACTTAAGCAAAACGATCCMYTWT 2398
|||

RESULT 31

US-09-962-436-279
; Sequence 279, Application US/09962436
; Patent No. US20020081301A1

GENERAL INFORMATION:

; APPLICANT: Soppet, Daniel
; TITLE OF INVENTION: Cancer Gene Determination and Therapeutic Screening Using Signal
; FILE REFERENCE: 689290-75
; CURRENT APPLICATION NUMBER: US/09/962,436
; CURRENT FILING DATE: 2001-09-25
; PRIOR APPLICATION NUMBER: US/60/235,082
; PRIOR FILING DATE: 2000-09-25
; PRIOR APPLICATION NUMBER: US/60/234,924
; PRIOR FILING DATE: 2000-09-25
; NUMBER OF SEQ ID NOS: 568
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 279
; LENGTH: 8210
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-962-436-279

Query Match 13.0%; Score 26.8; DB 10; Length 8210;
Best Local Similarity 52.7%; Pred. No. 34;
Matches 58; Conservative 0; Mismatches 52; Indels 0; Gaps 0;

QY 33 CAAGAACCAGTTAAAGGTCCTGTGTCTACTAAGCCAGGTTCTTGCTCCTATTATCTTGATT 92
|||
Db 6952 CATGAACCATTTCTTGACATTTGGCCTATATTGTCAGGGTTGATGATATGGTGGAAATT 7011
|||
QY 93 CGTTGCGCTATGTTAAACCCACCTAACCGTTGTTTGAAGGACACTGATTG 142
|||
Db 7012 CTTGCTTCTCTTCAGAAACCAAGAAAGATTCTTTAAAGGCTCAGATGG 7061
|||

RESULT 32

US-09-070-927A-612/c
; Sequence 612, Application US/09070927A
; Patent No. US20020120116A1

GENERAL INFORMATION:

; APPLICANT: Charles A. Kunsch
; Patrick J. Dillon
; Steven Barash
; TITLE OF INVENTION: Enterococcus faecialis Polynucleotides and Polypeptides
; NUMBER OF SEQUENCES: 982
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Human Genome Sciences, Inc.
; STREET: 9410 Key West Avenue
; CITY: Rockville

09-833799-13a.rnpb

```

; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; NUMBER OF SEQ ID NOS: 14110
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 7107
; LENGTH: 891
; TYPE: DNA
; ORGANISM: Haemophilus influenzae
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)...(891)
;
ts-09-815-242-7107

```

QY	154	AAAGTGCTGTGAAGGTTCTCGGGTATGGCTTGTTCGTT	194
D6	415	AATTGCAGTCAAGGTTTTGCTGTTTGTCTGTTCGTTCCGT	376

Query Match	12.9%;	Score 26.6;	DB 10;	Length 1857;
Best Local Similarity	58.0%;	Pred. No. 20;		
Matches 47: Conservative	0;	Mismatches	34;	Indels 0;
		Gaps	0;	

```

RESULT 35
US-09-070-927A-72
; Sequence 72, Application US/09070927A
; Patent No. US20020120116A1
; GENERAL INFORMATION:
; APPLICANT: Charles A. Kunsch
; Patrick J. Dillon
; Steven Barash
; TITLE OF INVENTION: Enterococci

```

RESULT 33
US-09-815-242-7107/c
; Sequence 7107, Application US/09815242
; Patent No. US20020061569A1
; GENERAL INFORMATION:
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari L.
; APPLICANT: Zyskind, Judith W.
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John D.
; APPLICANT: Carr, Grant J.
; APPLICANT: Yamamoto, Robert T.
; APPLICANT: Xu, H. Howard
; TITLE OF INVENTION: Identification of Essential Genes in
; TITLE OF INVENTION: Prokaryotes
; FILE REFERENCE: ELITRA.011A
; CURRENT APPLICATION NUMBER: US/09/815,242
; CURRENT FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/242,578

09-833799-13a.rnpb

Sun Feb 16 09:12:56 2003

FILE REFERENCE: SRIPI300-3
CURRENT APPLICATION NUMBER: US/09/938,842A
CURRENT FILING DATE: 2001-08-24
PRIOR APPLICATION NUMBER: US 60/227,866
PRIOR FILING DATE: 2000-08-24
PRIOR APPLICATION NUMBER: US 60/264,647
PRIOR FILING DATE: 2001-01-16
PRIOR APPLICATION NUMBER: US 60/300,111
PRIOR FILING DATE: 2001-06-22
NUMBER OF SEQ ID NOS: 5379
SEQ ID NO 4875
LENGTH: 2000
TYPE: DNA
ORGANISM: Arabidopsis thaliana
US-09-938-842A-4875

Query Match 12.7%; Score 26.2; DB 9; Length 2000;
Best Local Similarity 58.2%; Pred. No. 28;
Matches 46; Conservative 0; Mismatches 33; Indels 0; Gaps 0;

QY 31 CTCAAGAACCCAGTTAAAGGTCCTGTCTACTAAGCCAGGTTCTTGTCTTATTTCTGA 90
DB 843 CACGAAACATGTAATAAACTTCTAGTGTCAATAAAATATGTTATGTTAATAGTAA 784

QY 91 TTCGTTGCGCTATGTTAAA 109
DB 783 TCAGATGCAAAAGTTAAA 765

RESULT 40

US-09-822-849A-298/c
Sequence 298, Application US/09822849A
Patent No. US20020045170A1
GENERAL INFORMATION:
APPLICANT: Wong, Gordon G.
APPLICANT: Clark, Hilary
APPLICANT: Focht, Kim
APPLICANT: Agostino, Michael J.
APPLICANT: Howes, Steven H.
APPLICANT: Resnick, Richard J.
APPLICANT: Gulukota, Kamalakara
APPLICANT: Graham, James R.
APPLICANT: Genetics Institute, Inc.
TITLE OF INVENTION: POLYNUCLEOTIDES ENCODING NOVEL SECRETED PROTEINS
FILE REFERENCE: GIN 6403
CURRENT APPLICATION NUMBER: US/09/822,849A
CURRENT FILING DATE: 2001-09-04
PRIOR APPLICATION NUMBER: 60/195,582
PRIOR FILING DATE: 2000-04-06
NUMBER OF SEQ ID NOS: 598
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 298
LENGTH: 3479
TYPE: DNA
ORGANISM: Homo sapiens
US-09-822-849A-298

Query Match 12.7%; Score 26.2; DB 10; Length 3479;
Best Local Similarity 50.4%; Pred. No. 37;
Matches 64; Conservative 0; Mismatches 63; Indels 0; Gaps 0;

QY 1 AATTCGAGCTCGGTACCATACCTGCATATGCTCAAGAACCCAGTTAAAGGTCCTGTGCTA 60
DB 1948 AAATCGAGTTGTAAATAATGCTGCATAAATATGTTCCCATATTTGGTCTTTT 1889

QY 61 CTAAGCCAGGTTCTTGTCTTATTTATGTTGCTGCTATGTTAAACCCACCTAAC 120
DB 1888 AAGAGATGGGCTCGCTCTGTGCGCAGGCTGGAATGCATTAGTGCATACACAGTCACT 1829

QY 121 GTTGT 127
DB 1828 GCAGCTT 1822

NAME/KEY: SITE
LOCATION: (22)
OTHER INFORMATION: n equals a,t,g, or c
NAME/KEY: SITE
LOCATION: (24)
OTHER INFORMATION: n equals a,t,g, or c
US-09-764-869-355

Query Match 12.7%; Score 26.2; DB 10; Length 670;
Best Local Similarity 51.5%; Pred. No. 17;
Matches 52; Conservative 3; Mismatches 46; Indels 0; Gaps 0;

QY 90 ATTCGTTGCGCTATGTTAAACCCACCTAACCGTTTGTGAAGGACACTGATTGTCAGGT 149
DB 253 AATCTTAGGTTGGATAAAAGGCCCTTCTTTGTCTAAAAAGAGTAATCTCCATT 194

QY 150 ATCAAAAAGTGTGTGAAGGTTCTGCGGTATGGCTTGT 190
DB 193 ACKGTAAGGATTTCTAWAKGTTTCATGGCATCGGACWATTT 153

RESULT 38

US-10-105-930-1
Sequence 1, Application US/10105930
Publication No. US20030009018A1
GENERAL INFORMATION:
APPLICANT: Maeda, Masatsugu
APPLICANT: Yaguchi, No. US20030009018A1
TITLE OF INVENTION: NOVEL HEMOPOIETIN RECEPTOR PROTEIN, NR12
FILE REFERENCE: 06501-105US1
CURRENT APPLICATION NUMBER: US/10/105,930
CURRENT FILING DATE: 2002-03-25
PRIOR APPLICATION NUMBER: PCT/JP00/06654
PRIOR FILING DATE: 2000-09-27
PRIOR APPLICATION NUMBER: JP 2000-240397
PRIOR FILING DATE: 2000-08-03
PRIOR APPLICATION NUMBER: JP 11-273358
PRIOR FILING DATE: 1999-09-27
NUMBER OF SEQ ID NOS: 77
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 1
LENGTH: 1784
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: CDS
LOCATION: (98)...(1108)
US-10-105-930-1

Query Match 12.7%; Score 26.2; DB 9; Length 1784;
Best Local Similarity 52.3%; Pred. No. 27;
Matches 58; Conservative 0; Mismatches 53; Indels 0; Gaps 0;

QY 81 ATTATCTTGATTCGTTGCGCTATGTTAAACCCACCTAACCGTTGTTTGAAGGACACTGAT 140
DB 1321 AGTTTGATGCTGCAGTGAGCTATGATCATCCCACTGCTCTAGCTGAGGAAAGACCA 1380

QY 141 TGTCAGGATCAAAAAGTGTGTAAGGTTCTGCGGTATGGCTTGTTC 191
DB 1381 AGACCTGTTTCTTAAAGTTTAAACAGCCAGGTGCAGTGGCTTATGTC 1431

RESULT 39

US-09-938-842A-4875/c
Sequence 4875, Application US/09938842A
Patent No. US20020160378A1
GENERAL INFORMATION:
APPLICANT: Harper, Jeff
APPLICANT: Kreps, Joel
APPLICANT: Wang, Xun
APPLICANT: Zhu, Tong
TITLE OF INVENTION: STRESS-REGULATED GENES OF PLANTS, TRANSGENIC PLANTS CONTAINING
SAME, AND METHODS OF USE

RESULT 41

US-09-935-464-1/c
; Sequence 1, Application US/09935464
; Publication No. US20030027153A1
; GENERAL INFORMATION:
; APPLICANT: Meyer, Joanne
; APPLICANT: Barrington-Martin, Rory
; APPLICANT: Parker, Alexander
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR DIAGNOSING AND TREATING NEUROPSYCHIA
; TITLE OF INVENTION: DISORDERS SUCH AS SCHIZOPHRENIA
; FILE REFERENCE: 3322/1H702 US1
; CURRENT APPLICATION NUMBER: US/09/935,464
; CURRENT FILING DATE: 2001-08-23
; PRIOR APPLICATION NUMBER: US 09/757,300
; PRIOR FILING DATE: 2001-01-09
; NUMBER OF SEQ ID NOS: 90
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 1
; LENGTH: 157875
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-935-464-1

Query Match 12.7%; Score 26.2; DB 9; Length 157875;
Best Local Similarity 56.3%; Pred. No. 2.2e+02;
Matches 49; Conservative 0; Mismatches 38; Indels 0; Gaps 0;

Qy 61 CTAAGCCAGGTTCTTGTCTCTATATCTTGATTCGTGGCGCTATGTTAAACCCACCTAACC 120
Db 138995 CTCATCCTGCTCTTTGGCTTTTGGTCTTTAGGCTTTGGCGCTCTCTATAGTAGCTGAAGG 138936

QY 121 GTTGTTGAAGGACACTGATTGTCCAG 147

Db 138935 CCTGTAGGAATACACTCTTTGTACAG 138909

RESULT 42

US-09-790-988-1
; Sequence 1, Application US/09790988
; Patent No. US20020127687A1
; GENERAL INFORMATION:
; APPLICANT: SHIGENOBU, SHUJI
; APPLICANT: WATANABE, HIDEKI
; APPLICANT: HATTORI, MASAHIRA
; APPLICANT: SAKAKI, YOSHIYUKI
; TITLE OF INVENTION: GENOME DNA OF BACTERIAL SYMBIONT OF APHIDS
; FILE REFERENCE: 081356/0159
; CURRENT APPLICATION NUMBER: US/09/790,988
; CURRENT FILING DATE: 2001-02-23
; PRIOR APPLICATION NUMBER: JP2000-107160
; PRIOR FILING DATE: 2000-04-07
; NUMBER OF SEQ ID NOS: 7
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1
; LENGTH: 640681
; TYPE: DNA
; ORGANISM: Buchnera sp.
US-09-790-988-1

```
Query Match      12.7%; Score 26.2; DB 10; Length 640681;
Best Local Similarity 67.3%; Pred. No. 3.4e+02;
Matches 37; Conservative 0; Mismatches 18; Indels 0; Gaps 0;
```

Qy 62 TAAGCCAGGTTCTTGTCTATATATCTTGATTGGTTGCGCTATGTTAAACCCACCT 116
|||||
Dd 59718 TAAGCCAGTTTCTTCTTCTATTTTTTAATTAATTTTTTCACACTAATAAACACATCT 59772

RESULT 43

US-10-214-188-3
; Sequence 3, Application US/10214188
; Publication No. US20030022260A1

```

;
; GENERAL INFORMATION:
;
; APPLICANT: LA THANGUE, NICHOLAS B.
;           BERNARDS, RENE
;           HIJWANS, ELEANORE M.
;
; TITLE OF INVENTION: TRANSCRIPTION FACTOR E2F-5
;
; NUMBER OF SEQUENCES: 25
;
; CORRESPONDENCE ADDRESS:
;
; ADDRESSEE: NIXON & VANDERHYE P.C.
;
; STREET: 1100 NORTH GLEBE ROAD
;
; CITY: ARLINGTON
;
; STATE: VIRGINIA
;
; COUNTRY: U.S.A.
;
; ZIP: 22201-4714
;
; COMPUTER READABLE FORM:
;
; MEDIUM TYPE: Floppy disk
;
; COMPUTER: IBM PC compatible
;
; OPERATING SYSTEM: PC-DOS/MS-DOS
;
; SOFTWARE: PatentIn Release #1.0, Version #1.30
;
; CURRENT APPLICATION DATA:
;
; APPLICATION NUMBER: US/10/214,188
;
; FILING DATE: 08-Aug-2002
;
; CLASSIFICATION: <Unknown>
;
; PRIOR APPLICATION DATA:
;
; APPLICATION NUMBER: US/08/894,139
;
; FILING DATE: 13-AUG-1997
;
; ATTORNEY/AGENT INFORMATION:
;
; NAME: WILSON, MARY J.
;
; REGISTRATION NUMBER: 32,955
;
; REFERENCE/DOCKET NUMBER: 620-22
;
; TELECOMMUNICATION INFORMATION:
;
; TELEPHONE: (703) 816-4000
;
; TELEFAX: (703) 816-4100
;
; INFORMATION FOR SEQ ID NO: 3:
;
; SEQUENCE CHARACTERISTICS:
;
; LENGTH: 1340 base pairs
;
; TYPE: nucleic acid
;
; STRANDEDNESS: double
;
; TOPOLOGY: linear
;
; MOLECULE TYPE: DNA (genomic)
;
; FEATURE:
;
; NAME/KEY: CDS
;
; LOCATION: 16..1020
;
; SEQUENCE DESCRIPTION: SEQ ID NO: 3:
;

```

Query Match 12.6%; Score 26; DB 9; Length 1340;
Best Local Similarity 59.5%; Pred. No. 28;
Matches 44; Conservative 0; Mismatches 30; Indels

QY	68	AGTTCCTTGTCCTATTATCTTGATTTCGTTGCCTATGTTAAACCCACCTTAACCGTTGT	127
Db	1265	AGTTCCTCTCCTGAGAGGCAGTTACAGCAGGCTCAGTTCATCCAAAACAACATTGTCA	1324
QY	128	GAAGGACACTGATT	141
Db	1325	GAAGTACACTTATT	1338

RESULT 44

US-09-938-842A-4030/c
; Sequence 4030, Application US/09938842A
; Patent No. US20020160378A1
; GENERAL INFORMATION:
; APPLICANT: Harper, Jeff
; APPLICANT: Kreps, Joel
; APPLICANT: Wang, Xun
; APPLICANT: Zhu, Tong
; TITLE OF INVENTION: STRESS-REGULATED GENES OF PLANTS, TRANSGENIC PLANTS CONTAINING
; TITLE OF INVENTION: SAME, AND METHODS OF USE
; FILE REFERENCE: SCRI1300-3
; CURRENT APPLICATION NUMBER: US/09/938,842A
; CURRENT FILING DATE: 2001-08-24
; PRIOR APPLICATION NUMBER: US 60/227,866

Sun Feb 16 09:12:56 2003

ZIP: 20850
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette, 3.50 inch, 1.4Mb storage
COMPUTER: HP Vectra 486/33
OPERATING SYSTEM: MSDOS version 6.2
SOFTWARE: ASCII Text
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/070,927A
FILING DATE: 04-May-2000
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 60/046,655
FILING DATE: 1997-05-16
APPLICATION NUMBER: 60/044,031
FILING DATE: 1997-05-06
APPLICATION NUMBER: 60/066,009
FILING DATE: 1997-11-14
ATTORNEY/AGENT INFORMATION:
NAME: Kenley K. Hoover
REGISTRATION NUMBER: 40,302
REFERENCE/DOCKET NUMBER: PB369
TELECOMMUNICATION INFORMATION:
TELEPHONE: (301) 309-8504
TELEFAX: (301) 309-8512
INFORMATION FOR SEQ ID NO: 722:
SEQUENCE CHARACTERISTICS:
LENGTH: 539 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
SEQUENCE DESCRIPTION: SEQ ID NO: 722:
US-09-070-927A-722

Query Match 12.6%; Score 26; DB 9; Length 2000;
Best Local Similarity 59.5%; Pred. No. 33;
Matches 44; Conservative 0; Mismatches 30; Indels 0; Gaps 0;

QY 20 ACCTGCATATGCTCAAGAACCAAGTTAAAGGTCCTGTCTACTAAGCCAGGTTCTTGTC 79
DB 144 AACTGATTATGATTATAAGTAATAAAATTAGTAATGTTTACGTTACCAAGTTCTTTTCA 85
QY 80 TATTATCTTGATTC 93
DB 84 TTTTCTTTTTC 71

RESULT 45
US-09-764-877-2223
Sequence 2223, Application US/09764877
Patent No. US20020147140A1
GENERAL INFORMATION:
APPLICANT: Rosen et al.
TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
FILE REFERENCE: PC005
CURRENT APPLICATION NUMBER: US/09/764,877
CURRENT FILING DATE: 2001-01-17
Prior application data removed - refer to PALM or file wrapper
NUMBER OF SEQ ID NOS: 4031
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 2223
LENGTH: 340
TYPE: DNA
ORGANISM: Homo sapiens
US-09-764-877-2223

Query Match 12.5%; Score 25.8; DB 10; Length 340;
Best Local Similarity 60.9%; Pred. No. 17;
Matches 42; Conservative 0; Mismatches 27; Indels 0; Gaps 0;

QY 89 GATTCGTTGCGTATGTTAAACCCACCTAACCCTGTTTGAAGGACACTGATGTCACG 148
DB 261 GTTTCAGTGAGCCCAAGTTTCATGCCCGCACTCTAGTTTGGCGACAGCAAGACTCTG 320
QY 149 TATCAAAA 157
DB 321 TCTCAAAA 329

RESULT 46
US-09-070-927A-722
Sequence 722, Application US/09070927A
Patent No. US20020120116A1
GENERAL INFORMATION:
APPLICANT: Charles A. Kunsch
Steven Barash
Patrick J. Dillon
TITLE OF INVENTION: Enterococcus faecialis Polynucleotides and Polypeptides
NUMBER OF SEQUENCES: 982
CORRESPONDENCE ADDRESS:
ADDRESSEE: Human Genome Sciences, Inc.
STREET: 9410 Key West Avenue
CITY: Rockville
STATE: Maryland
COUNTRY: USA

Query Match 12.5%; Score 25.8; DB 10; Length 539;
Best Local Similarity 48.9%; Pred. No. 21;
Matches 66; Conservative 0; Mismatches 69; Indels 0; Gaps 0;

QY 57 TCTACTAAGCCAGGTTCTTCTCTATTATCTTGAATTCGTCGCTATGTTAAACCCACCT 116
DB 28 TCTAACATGTCTGGTNCCTTTCTTAACATCTTAGCTAGAGCCTTAGGTNTCTTTCACAC 87
QY 117 AACCGTTGTTGAAGGACACTGATGTCAGGTATCAAAAAGTGTGTGAAGGTTCTGTC 176
DB 88 TTCCTCTTCCCTAAGTATTTCTTTGTAACCGTCATGGAAGAAAGTAATCTCGCTTTCAC 147
QY 177 GGTATGGCTTGTTC 191
DB 148 GTTAAAAATTGGATC 162

RESULT 47
US-09-886-055-410
Sequence 410, Application US/09886055
Patent No. US20020132273A1
GENERAL INFORMATION:
APPLICANT: STRYER, LUBERT
APPLICANT: ZOZULYA, SERGEY
TITLE OF INVENTION: RECEPTOR FINGERPRINTING, SENSORY PERCEPTION, AND
TITLE OF INVENTION: BIOSENSORS OF CHEMICAL SENSANTS
FILE REFERENCE: 078003-0277150
CURRENT APPLICATION NUMBER: US/09/886,055
CURRENT FILING DATE: 2001-06-22
PRIOR APPLICATION NUMBER: 60/213,812
PRIOR FILING DATE: 2000-06-22
NUMBER OF SEQ ID NOS: 522
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 410
LENGTH: 1047
TYPE: DNA
ORGANISM: Homo sapiens
US-09-886-055-410

Query Match 12.5%; Score 25.8; DB 10; Length 1047;

```

; SOFTWARE: PatentIn ver. 3.0
; SEQ ID NO 1385
; LENGTH: 1107
; TYPE: DNA
; ORGANISM: Corynebacterium glutamicum
US-09-738-626-1385

Query Match          12.5%;   Score 25.8;   DB 9;   Length 1107;
Best Local Similarity 58.4%;   Pred. No. 30;
Matches 45;   Conservative 0;   Mismatches 32;   Indels 0;   Gaps 0;

```

QY	50	TCCTGTGCTACTAAGCCAGGTTCTTGTCCTATTATCTTGAATCGTTGCGCTATGTTAAA	109
Dδ	447	TCCCCTGTTAAACCAAGTCAGTGACTTGTTCCCTCTGACTGAATGTTGTGGCACCCAGATGAT	388
QY	110	CCCACCTAACCGTTGTT	126
Dδ	387	TCCATCGAGGCGTTGGT	371

```

RESULT 50
US-09-965-528-31
; Sequence 31, Application US/09965528
; Publication No. US20020187523A1
; GENERAL INFORMATION:
; APPLICANT: INCYTE GENOMICS, INC.
; APPLICANT: TANG, Y. Tom
; APPLICANT: YUE, Henry
; APPLICANT: LAL, Preeti
; APPLICANT: BURFORD, Neil
; APPLICANT: BANDMAN, Olga
; APPLICANT: BAUGHN, Mariah R.
; APPLICANT: AZIMZAI, Yalda
; APPLICANT: LU, Dyung Aina M.
; APPLICANT: PATTERSON, Chandra
; TITLE OF INVENTION: EXTRACELLULAR SIGNALING MOLECULES
; FILE REFERENCE: PF-0701 USA
; CURRENT APPLICATION NUMBER: US/09/965,528
; CURRENT FILING DATE: 2001-09-26
; PRIOR APPLICATION NUMBER: 60/134,949
; PRIOR FILING DATE: 1999-05-19
; PRIOR APPLICATION NUMBER: 60/144,270
; PRIOR FILING DATE: 1999-07-15
; PRIOR APPLICATION NUMBER: 60/146,700
; PRIOR FILING DATE: 1999-07-30
; PRIOR APPLICATION NUMBER: 60/157,508
; PRIOR FILING DATE: 1999-10-04
; NUMBER OF SEQ ID NOS: 55
; SOFTWARE: PERL Program
; SEQ ID NO 31
; LENGTH: 1125
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; OTHER INFORMATION: Incyte ID No. US20020187523A1 156
US-09-965-528-31

```

	Query Match	12.5%;	Score 25.8;	DB 9;	Length 1125;
	Best Local Similarity	56.5%;	Pred. NO. 30;		
	Matches 48; Conservative	0;	Mismatches 37;	Indels 0;	Gaps 0;
QY	84	ATCTTGATTGCTTGGCGTATGTATAACCCACCTAACCGTTGTTTGAAGGACACTGATTGT	143		
Dd	527	ATTTTGAAGTGGTTGAAGAGGAGAACTTTTCCTAACAACCTGTTCAAAGACTCTTTTACT	586		
QY	144	CCAGGTATCAAAAAAGTGTGTGAAG	168		
Dd	587	CCAGGGAACATAGCTAACTGGGAAG	611		

RESULT 51
US-09-938-540-1/c

Sun Feb 16 09:12:56 2003

GENERAL INFORMATION:
; APPLICANT: Olga Bandman
; TITLE OF INVENTION: GENES DIFFERENTIALLY EXPRESSED IN VASCULAR TISSUE ACTIVATION
; FILE REFERENCE: PA-0028 US
; CURRENT APPLICATION NUMBER: US/10/044,090
; CURRENT FILING DATE: 2002-01-09
; NUMBER OF SEQ ID NOS: 850
; SOFTWARE: PERL Program
; SEQ ID NO 333
; LENGTH: 3189
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; OTHER INFORMATION: Incyte ID No. US20020137081A1 2960079CB1
US-10-044-090-333

Query Match 12.5%; Score 25.8; DB 12; Length 3189;
Best Local Similarity 52.3%; Pred. No. 49;
Matches 57; Conservative 0; Mismatches 52; Indels 0; Gaps 0;

QY 71 TTCTTGTCTTATTTATCTTGTGATTCGTCGCTATGTTAAACCCACCTAACCGTTGTTTGA 130
Db 2743 TTCTTTCCTCTGATCTGAATCAGCTGAGCTGCTGAGCAGACATATTACTGCTGTGGATA 2684
QY 131 GGACACTGATTGTCAGGTATCAAAAAGTCTGTGAAGTTCTCTGCGGT 179
Db 2683 GTAAGACTGCTGTGGGGCTGAGGGAAGGGGTATGAAGGCTGCTGGGGT 2635

RESULT 54
US-09-800-729-68/c
; Sequence 68, Application US/09800729
; Patent No. US20020068319A1
; GENERAL INFORMATION:
; APPLICANT: Ni et al.
; TITLE OF INVENTION: 32 Human secreted proteins
; FILE REFERENCE: PZ044P1
; CURRENT APPLICATION NUMBER: US/09/800,729
; CURRENT FILING DATE: 2001-03-08
; PRIOR APPLICATION NUMBER: PCT/US00/26013
; PRIOR FILING DATE: 2000-09-22
; PRIOR APPLICATION NUMBER: 60/155,709
; PRIOR FILING DATE: 1999-09-24
; NUMBER OF SEQ ID NOS: 217
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 68
; LENGTH: 3299
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-800-729-68

Query Match 12.5%; Score 25.8; DB 10; Length 3299;
Best Local Similarity 52.3%; Pred. No. 50;
Matches 57; Conservative 0; Mismatches 52; Indels 0; Gaps 0;

QY 71 TTCTTGTCTTATTTATCTTGTGATTCGTTGCGCTATGTTAAACCCACCTAACCGTTGTTTGA 130
Db 2813 TTCTTTCCTCTGATCTGAATCAGCTGAGCTGCTGAGCAGACATATTACTGCTGTGGATA 2754
QY 131 GGACACTGATTGTCAGGTATCAAAAAGTCTGTGAAGTTCTCTGCGGT 179
Db 2753 GTAAGACTGCTGTGGGGCTGAGGGAAGGGGTATGAAGGCTGCTGGGGT 2705

RESULT 55
US-09-822-846-104/c
; Sequence 104, Application US/09822846
; Publication No. US20030027139A1
; GENERAL INFORMATION:
; APPLICANT: Jacobs, Kenneth
; APPLICANT: McCoy, John M.
; APPLICANT: Lavallie, Edward R.

Sequence 1, Application US/09938540
; Patent No. US20020151001A1
; GENERAL INFORMATION:
; APPLICANT: Degussa AG
; TITLE OF INVENTION: New nucleotide sequences which code for the ccpA1 gene
; FILE REFERENCE: 000059 BT
; CURRENT APPLICATION NUMBER: US/09/938,540
; CURRENT FILING DATE: 2001-08-27
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1
; LENGTH: 1600
; TYPE: DNA
; ORGANISM: Corynebacterium glutamicum
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (225)..(1388)
; OTHER INFORMATION: ccpA1-Gen
US-09-938-540-1

Query Match 12.5%; Score 25.8; DB 10; Length 1600;
Best Local Similarity 58.4%; Pred. No. 35;
Matches 45; Conservative 0; Mismatches 32; Indels 0; Gaps 0;

QY 50 TCCTGTGCTACTAAGCCAGGTTCTTGTCTTATTTATCTTGTGATTCGTTGCGCTATGTTTAA 109
Db 728 TCCCTGTTAAACCAAGTCAGTGACTTGTTCCTCTGACTGAATGTGTGGCACCACGATGAT 669
QY 110 CCCACCTAACCGTTGTT 126
Db 668 TCCATCGAGCGTTGGT 652

RESULT 52
US-09-925-300-693/c
; Sequence 693, Application US/09925300
; Patent No. US20020151681A1
; GENERAL INFORMATION:
; APPLICANT: Craig Rosen,
; APPLICANT: Steve Ruben
; TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies
; FILE REFERENCE: PA101
; CURRENT APPLICATION NUMBER: US/09/925,300
; CURRENT FILING DATE: 2001-08-10
; PRIOR APPLICATION NUMBER: PCT/US00/05988
; PRIOR FILING DATE: 2000-03-08
; PRIOR APPLICATION NUMBER: 60/124,270
; PRIOR FILING DATE: 1999-03-12
; NUMBER OF SEQ ID NOS: 1890
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 693
; LENGTH: 3098
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-925-300-693

Query Match 12.5%; Score 25.8; DB 10; Length 3098;
Best Local Similarity 52.3%; Pred. No. 48;
Matches 57; Conservative 0; Mismatches 52; Indels 0; Gaps 0;

QY 71 TTCTTGTCTTATTTATCTTGTGATTCGTTGCGCTATGTTAAACCCACCTAACCGTTGTTTGA 130
Db 2612 TTCTTTCCTCTGATCTGAATCAGCTGAGCTGCTGAGCAGACATATTACTGCTGTGGATA 2553
QY 131 GGACACTGATTGTCAGGTATCAAAAAGTCTGTGAAGTTCTCTGCGGT 179
Db 2552 GTAAGACTGCTGTGGGGCTGAGGGAAGGGGTATGAAGGCTGCTGGGGT 2504

RESULT 53
US-10-044-090-333/c
; Sequence 333, Application US/10044090
; Patent No. US20020137081A1

APPLICANT: Collins-Racie, Lisa A.
APPLICANT: Evans, Cheryl
APPLICANT: Merberg, David
APPLICANT: Treacy, Maurice
APPLICANT: Agostino, Michael J.
APPLICANT: Steininger II, Robert J.
APPLICANT: Bowman, Michael R.
APPLICANT: Spaulding, Vikki
APPLICANT: Wong, Gordon G.
APPLICANT: Clark, Hilary
APPLICANT: Fechtel, Kim
APPLICANT: Howes, Steven H.
APPLICANT: Resnick, Richard J.
APPLICANT: Gulukota, Kamalakara
APPLICANT: Graham, James R.
APPLICANT: Genetics Institute, Inc.
TITLE OF INVENTION: POLYNUCLEOTIDES ENCODING NOVEL SECRETED PROTEINS
FILE REFERENCE: GIN 6400
CURRENT APPLICATION NUMBER: US/09/822,846
CURRENT FILING DATE: 2001-03-29
PRIOR APPLICATION NUMBER: 60/195,605
PRIOR FILING DATE: 2000-04-06
NUMBER OF SEQ ID NOS: 629
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 104
LENGTH: 4116
TYPE: DNA
ORGANISM: Homo sapiens
US-09-822-846-104

Query Match 12.5%; Score 25.8; DB 9; Length 4116;
Best Local Similarity 52.3%; Pred. No. 55;
Matches 57; Conservative 0; Mismatches 52; Indels 0; Gaps 0;

QY 71 TTCTTGCTCTATTATCTTGATTTCGTTGGCTATGTTAAACCCACCTAACCGTTGTTGAA 130
|||||
Db 977 TTCTTCCCTCTGATCTGAATCAGCTGAGCTGCTGAGCAGACATATTACTGCTGTGGATA 918
|||||

QY 131 GGACACTGATTGTCAGGTATCAAAAAGTGTGTGAAGTTCTCTGCGGT 179
|||
Db 917 GTAAGACTGCTGTGGGGCTGAGGGAAGGGGTATGAAGGCTGCTGGGGT 869
|||

RESULT 56
US-09-800-729-35/c
; Sequence 35, Application US/09800729
; Patent No. US20020068319A1
; GENERAL INFORMATION:
; APPLICANT: Ni et al.
; TITLE OF INVENTION: 32 Human secreted proteins
; FILE REFERENCE: PZ044P1
; CURRENT APPLICATION NUMBER: US/09/800,729
; CURRENT FILING DATE: 2001-03-08
; PRIOR APPLICATION NUMBER: PCT/US00/26013
; PRIOR FILING DATE: 2000-09-22
; PRIOR APPLICATION NUMBER: 60/155,709
; PRIOR FILING DATE: 1999-09-24
; NUMBER OF SEQ ID NOS: 217
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 35
; LENGTH: 6065
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (6035)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: SITE
; LOCATION: (6037)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: SITE
; LOCATION: (6038)
; OTHER INFORMATION: n equals a,t,g, or c

US-09-800-729-35
Query Match 12.5%; Score 25.8; DB 10; Length 6065;
Best Local Similarity 52.3%; Pred. No. 67;
Matches 57; Conservative 0; Mismatches 52; Indels 0; Gaps 0;

QY 71 TTCTTGCTCTATTATCTTGATTTCGTTGGCTATGTTAAACCCACCTAACCGTTGTTGAA 130
|||||
Db 2813 TTCTTCCCTCTGATCTGAATCAGCTGAGCTGCTGAGCAGACATATTACTGCTGTGGATA 2754
|||||

QY 131 GGACACTGATTGTCAGGTATCAAAAAGTGTGTGAAGTTCTCTGCGGT 179
|||
Db 2753 GTAAGACTGCTGTGGGGCTGAGGGAAGGGGTATGAAGGCTGCTGGGGT 2705
|||||

RESULT 57
US-10-002-048A-4
; Sequence 4, Application US/10002048A
; Publication No. US20020182616A1
; GENERAL INFORMATION:
; APPLICANT: Wahlestedt, Claes
; APPLICANT: Ding, Bo
; TITLE OF INVENTION: Single Nucleotide Polymorphisms
; FILE REFERENCE: 10806-143
; CURRENT APPLICATION NUMBER: US/10/002,048A
; CURRENT FILING DATE: 2001-11-02
; PRIOR APPLICATION NUMBER: SE 0004035-2
; PRIOR FILING DATE: 2000-11-03
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 4
; LENGTH: 247
; TYPE: DNA
; ORGANISM: human
; FEATURE:
; NAME/KEY: Intron
; LOCATION: (1)..(29)
; OTHER INFORMATION:
; FEATURE:
; NAME/KEY: Intron
; LOCATION: (218)..(247)
; OTHER INFORMATION:
; PUBLICATION INFORMATION:
; DATABASE ACCESSION NUMBER: GenBank/M14296
; DATABASE ENTRY DATE: 1995-01-08
US-10-002-048A-4

Query Match 12.4%; Score 25.6; DB 9; Length 247;
Best Local Similarity 55.7%; Pred. No. 17;
Matches 49; Conservative 0; Mismatches 39; Indels 0; Gaps 0;

QY 14 TACCATACCTGCATATGCTCAAGAACAGTTAAAGGTCCTGTCTACTAAGCCAGGTTTC 73
|||
Db 16 TTCTGTGCTGCAGATGCTAGGTAACAAGGACTGGGGCTGTCGGACTGACCCCTCGCCC 75
|||

QY 74 TTGTCTCTATTATCTTGATTTCGTTGCGCT 101
|||
Db 76 TGTCCCTGCTGCTGTCCTGGGTGCGCT 103
|||

RESULT 58
US-09-960-352-6148
; Sequence 6148, Application US/09960352
; Patent No. US20020137139A1
; GENERAL INFORMATION:
; APPLICANT: Warren, Wesley C.
; APPLICANT: Tao, Nengbing
; APPLICANT: Byatt, John C.
; APPLICANT: Mathialagan, Nagappan
; TITLE OF INVENTION: NUCLEIC ACID AND OTHER MOLECULES ASSOCIATED WITH LACTATION AND
; FILE REFERENCE: 16511.006/37-21(10298)C
; CURRENT APPLICATION NUMBER: US/09/960,352

09-833799-13a.rnpb

Sun Feb 16 09:12:56 2003

CURRENT FILING DATE: 2001-09-24
NUMBER OF SEQ ID NOS: 15112
SEQ ID NO 6148
LENGTH: 327
TYPE: DNA
ORGANISM: Bos taurus
OTHER INFORMATION: Clone ID: 26-LIB34-075-Q1-E1-G5
US-09-960-352-6148

Query Match
Best Local Similarity 12.4%; Score 25.6; DB 10; Length 327;
Matches 58; Conservative 0; Mismatches 54; Indels 0; Gaps 0;

QY 16 CCATACCTGCATATGCTCAAGAACACAGTTAAAGTCTGTCTACTAAGCCAGTTCTT 75
Db 160 CCACAACAGCATAAAGCTCAACGCACAGAGGATTGTGCATTGCATATTAGTCAAGATGAT 219
QY 76 GTCCTATTATCTTATTCGTTGCGGTATGTTAAACCCACCTAACCGTTGTTT 127
Db 220 TGCATCTCTTATTTCATCGACTTCCCAATGAAACATTCTGCTACTGTAGCTT 271

RESULT 59
US-09-954-456-2208/c
Sequence 2208, Application US/09954456
Patent No. US20020115057A1
GENERAL INFORMATION:
APPLICANT: Young, Paul
TITLE OF INVENTION: Process for Identifying Anti-Cancer Therapeutic Agents Using Cand
TITLE OF INVENTION: Sets
FILE REFERENCE: 689290-76
CURRENT APPLICATION NUMBER: US/09/954,456
CURRENT FILING DATE: 2001-09-18
PRIOR APPLICATION NUMBER: US/60/233,617
PRIOR FILING DATE: 2000-09-18
PRIOR APPLICATION NUMBER: US/60/234,052
PRIOR FILING DATE: 2000-09-20
PRIOR APPLICATION NUMBER: US/60/234,923
PRIOR FILING DATE: 2000-09-25
PRIOR APPLICATION NUMBER: US/60/235,134
PRIOR FILING DATE: 2000-09-25
PRIOR APPLICATION NUMBER: US/60/235,637
PRIOR FILING DATE: 2000-09-26
PRIOR APPLICATION NUMBER: US/60/235,638
PRIOR FILING DATE: 2000-09-26
PRIOR APPLICATION NUMBER: US/60/235,711
PRIOR FILING DATE: 2000-09-27
PRIOR APPLICATION NUMBER: US/60/235,720
PRIOR FILING DATE: 2000-09-27
PRIOR APPLICATION NUMBER: US/60/235,840
PRIOR FILING DATE: 2000-09-27
PRIOR APPLICATION NUMBER: US/60/235,863
PRIOR FILING DATE: 2000-09-27
NUMBER OF SEQ ID NOS: 2276
SOFTWARE: PatentIn version 3.0
SEQ ID NO 2208
LENGTH: 434
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: misc feature
OTHER INFORMATION: n=a,t,g or c
US-09-954-456-2208

Query Match
Best Local Similarity 12.4%; Score 25.6; DB 10; Length 434;
Matches 40; Conservative 0; Mismatches 24; Indels 0; Gaps 0;

QY 131 GGACACTGATTCAGGTATCAAAAAGTGTGTGAAGTTCTCGCGGTATGGCTTGT 190
Db 214 GAACTGTGGTGGCAAGGTATTTAAAGGGCTTTCTCTGCCTCTCTCTTTGATTATT 155
QY 191 CGTT 194

US-09-954-456-2208
Sequence 2208, Application US/09954456
Patent No. US20020115057A1
GENERAL INFORMATION:
APPLICANT: Young, Paul
TITLE OF INVENTION: Process for Identifying Anti-Cancer Therapeutic Agents Using Cand
TITLE OF INVENTION: Sets
FILE REFERENCE: 689290-76
CURRENT APPLICATION NUMBER: US/09/954,456
CURRENT FILING DATE: 2001-09-18
PRIOR APPLICATION NUMBER: US/60/233,617
PRIOR FILING DATE: 2000-09-18
PRIOR APPLICATION NUMBER: US/60/234,052
PRIOR FILING DATE: 2000-09-20
PRIOR APPLICATION NUMBER: US/60/234,923
PRIOR FILING DATE: 2000-09-25
PRIOR APPLICATION NUMBER: US/60/235,134
PRIOR FILING DATE: 2000-09-25
PRIOR APPLICATION NUMBER: US/60/235,637
PRIOR FILING DATE: 2000-09-26
PRIOR APPLICATION NUMBER: US/60/235,638
PRIOR FILING DATE: 2000-09-26
PRIOR APPLICATION NUMBER: US/60/235,711
PRIOR FILING DATE: 2000-09-27
PRIOR APPLICATION NUMBER: US/60/235,720
PRIOR FILING DATE: 2000-09-27
PRIOR APPLICATION NUMBER: US/60/235,840
PRIOR FILING DATE: 2000-09-27
PRIOR APPLICATION NUMBER: US/60/235,863
PRIOR FILING DATE: 2000-09-27
NUMBER OF SEQ ID NOS: 2276
SOFTWARE: PatentIn version 3.0
SEQ ID NO 2208
LENGTH: 434
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: misc feature
OTHER INFORMATION: n=a,t,g or c
US-09-954-456-2208

Query Match
Best Local Similarity 12.4%; Score 25.6; DB 10; Length 434;
Matches 40; Conservative 0; Mismatches 24; Indels 0; Gaps 0;

QY 131 GGACACTGATTCAGGTATCAAAAAGTGTGTGAAGTTCTCGCGGTATGGCTTGT 190
Db 214 GAACTGTGGTGGCAAGGTATTTAAAGGGCTTTCTCTGCCTCTCTCTTTGATTATT 155
QY 191 CGTT 194

Db 154 AATT 151

RESULT 60
US-09-880-107-2920/c
Sequence 2920, Application US/09880107
Patent No. US20020142981A1
GENERAL INFORMATION:
APPLICANT: Horne, Darci T.
APPLICANT: Vockley, Joseph G.
APPLICANT: Scherf, Uwe
APPLICANT: Gene Logic, Inc.
TITLE OF INVENTION: Gene Expression Profiles in Liver Cancer
FILE REFERENCE: 44921-5028-WO
CURRENT APPLICATION NUMBER: US/09/880,107
CURRENT FILING DATE: 2001-06-14
PRIOR APPLICATION NUMBER: US 60/211,379
PRIOR FILING DATE: 2000-06-14
PRIOR APPLICATION NUMBER: US 60/237,054
PRIOR FILING DATE: 2000-10-02
NUMBER OF SEQ ID NOS: 3950
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 2920
LENGTH: 434
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
OTHER INFORMATION: Genbank Accession No. US20020142981A1 R61374
NAME/KEY: unsure
LOCATION: (1)..(434)
OTHER INFORMATION: n = a or c or g or t
US-09-880-107-2920

Query Match
Best Local Similarity 12.4%; Score 25.6; DB 10; Length 434;
Matches 40; Conservative 0; Mismatches 24; Indels 0; Gaps 0;

QY 131 GGACACTGATTCAGGTATCAAAAAGTGTGTGAAGTTCTCGCGGTATGGCTTGT 190
Db 214 GAACTGTGGTGGCAAGGTATTTAAAGGGCTTTCTCTGCCTCTCTCTTTGATTATT 155
QY 191 CGTT 194
Db 154 AATT 151

RESULT 61
US-09-783-590-11050
Sequence 11050, Application US/09783590
Patent No. US20020110850A1
GENERAL INFORMATION:
APPLICANT: Dillon, Patrick J.
APPLICANT: Haseitine, William A.
APPLICANT: Li, Haodong
APPLICANT: Rosen, Craig A.
APPLICANT: Ruben, Steven M.
TITLE OF INVENTION: Human Genes, Sequences, and Expression Products 16.2
FILE REFERENCE: PO-16.2C1
CURRENT APPLICATION NUMBER: US/09/783,590
CURRENT FILING DATE: 2000-02-15
PRIOR APPLICATION NUMBER: 08/420,856
PRIOR FILING DATE: 1995-04-12
PRIOR APPLICATION NUMBER: 08/346,731
PRIOR FILING DATE: 1994-11-21
NUMBER OF SEQ ID NOS: 12485
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 11050
LENGTH: 468
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: misc feature

Sun Feb 16 09:12:56 2003

Db 279 TCTCGGAGTAGCTCGTTAGTTTCGGGTAGTCAAACTGAATTCGTTTGGCTTGATTATG 220
Qy 135 ACTGATTGTCAGGTATCAAAAAGTGTGTGAAGTTCTCGGGTATGGCTTGTTCGTT 194
Db 219 ACTGGCTAGACCATATGCAAAAGGTACGAGGTAAGTTCTCTGCTTTTGTGCTTAGTT 160

RESULT 64
US-09-815-242-8813
; Sequence 8813, Application US/09815242
; Patent No. US20020061569A1
; GENERAL INFORMATION:
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari L.
; APPLICANT: Zyskind, Judith W.
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John D.
; APPLICANT: Carr, Grant J.
; APPLICANT: Yamamoto, Robert T.
; APPLICANT: Xu, H. Howard
; TITLE OF INVENTION: Identification of Essential Genes in
; TITLE OF INVENTION: Prokaryotes
; FILE REFERENCE: ELITRA.011A
; CURRENT APPLICATION NUMBER: US/09/815,242
; CURRENT FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; NUMBER OF SEQ ID NOS: 14110
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 8813
; LENGTH: 942
; TYPE: DNA
; ORGANISM: Staphylococcus aureus
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)...(942)

US-09-815-242-8813
Query Match 12.4%; Score 25.6; DB 10; Length 942;
Best Local Similarity 59.7%; Pred. No. 32;
Matches 43; Conservative 0; Mismatches 29; Indels 0; Gaps 0;
Qy 101 TATGTTAAACCCACCTAACCGTTGTTTGAAGGACACTGATTGTCCAGGTATCAAAAAGTG 160
Db 806 TATGCTATATCGATAATATCGATGCTAGATGAATATGTTTGAAAAGGCATATAAAAAA 865
Qy 161 CTGTGAAGGTTTC 172
Db 866 CTGACAAGGGTTC 877

RESULT 65
US-09-815-242-8969
; Sequence 8969, Application US/09815242
; Patent No. US20020061569A1
; GENERAL INFORMATION:
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari L.
; APPLICANT: Zyskind, Judith W.
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John D.

US-09-815-242-8813
Query Match 12.4%; Score 25.6; DB 10; Length 942;
Best Local Similarity 59.7%; Pred. No. 32;
Matches 43; Conservative 0; Mismatches 29; Indels 0; Gaps 0;
Qy 101 TATGTTAAACCCACCTAACCGTTGTTTGAAGGACACTGATTGTCCAGGTATCAAAAAGTG 160
Db 806 TATGCTATATCGATAATATCGATGCTAGATGAATATGTTTGAAAAGGCATATAAAAAA 865
Qy 161 CTGTGAAGGTTTC 172
Db 866 CTGACAAGGGTTC 877

RESULT 66
US-09-995-898A-3
; Sequence 3, Application US/09995898A
; Publication No. US20030027253A1
; GENERAL INFORMATION:
; APPLICANT: Presnell, Scott R.
; APPLICANT: Xu, Wenfeng
; APPLICANT: No. US20030027253A1ak, Julia E.
; APPLICANT: Whitmore, Theodore E.
; APPLICANT: Grant, Francis J.
; TITLE OF INVENTION: CYTOKINE RECEPTOR ZCYTOR19
; FILE REFERENCE: 00-108
; CURRENT APPLICATION NUMBER: US/09/995,898A
; CURRENT FILING DATE: 2001-11-28
; PRIOR APPLICATION NUMBER: US 60/253,561
; PRIOR FILING DATE: 2000-11-28
; PRIOR APPLICATION NUMBER: US 60/267,211
; PRIOR FILING DATE: 2001-02-07
; NUMBER OF SEQ ID NOS: 50
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 3
; LENGTH: 1473
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Degenerate polynucleotide sequence of SEQ ID NO:2
; FEATURE:

; APPLICANT: Carr, Grant J.
; APPLICANT: Yamamoto, Robert T.
; APPLICANT: Xu, H. Howard
; TITLE OF INVENTION: Identification of Essential Genes in
; TITLE OF INVENTION: Prokaryotes
; FILE REFERENCE: ELITRA.011A
; CURRENT APPLICATION NUMBER: US/09/815,242
; CURRENT FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; NUMBER OF SEQ ID NOS: 14110
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 8969
; LENGTH: 942
; TYPE: DNA
; ORGANISM: Staphylococcus aureus
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)...(942)

US-09-815-242-8969
Query Match 12.4%; Score 25.6; DB 10; Length 942;
Best Local Similarity 59.7%; Pred. No. 32;
Matches 43; Conservative 0; Mismatches 29; Indels 0; Gaps 0;
Qy 101 TATGTTAAACCCACCTAACCGTTGTTTGAAGGACACTGATTGTCCAGGTATCAAAAAGTG 160
Db 806 TATGCTATATCGATAATATCGATGCTAGATGAATATGTTTGAAAAGGCATATAAAAAA 865
Qy 161 CTGTGAAGGTTTC 172
Db 866 CTGACAAGGGTTC 877

RESULT 66
US-09-995-898A-3
; Sequence 3, Application US/09995898A
; Publication No. US20030027253A1
; GENERAL INFORMATION:
; APPLICANT: Presnell, Scott R.
; APPLICANT: Xu, Wenfeng
; APPLICANT: No. US20030027253A1ak, Julia E.
; APPLICANT: Whitmore, Theodore E.
; APPLICANT: Grant, Francis J.
; TITLE OF INVENTION: CYTOKINE RECEPTOR ZCYTOR19
; FILE REFERENCE: 00-108
; CURRENT APPLICATION NUMBER: US/09/995,898A
; CURRENT FILING DATE: 2001-11-28
; PRIOR APPLICATION NUMBER: US 60/253,561
; PRIOR FILING DATE: 2000-11-28
; PRIOR APPLICATION NUMBER: US 60/267,211
; PRIOR FILING DATE: 2001-02-07
; NUMBER OF SEQ ID NOS: 50
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 3
; LENGTH: 1473
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Degenerate polynucleotide sequence of SEQ ID NO:2
; FEATURE:

```

; NAME/KEY: misc feature
; LOCATION: (1)...(1473)
; OTHER INFORMATION: n = A,T,C or G
US-09-995-898A-3

Query Match      12.4%; Score 25.6; DB 9; Length 1473;
Best Local Similarity 28.9%; Pred. No. 40;
Matches 37; Conservative 24; Mismatches 67; Indels 0; Gaps 0;

QY 1 AATTCGAGCTCGGTACCATACCTGCATATGCTCAAGAACCAAGTTAAAGTCCCTGTCTA 60
   :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :|
Db 264 RAARCARGAYTNTAYAAAYARTTYAARGNGMGTNMGACNGTNSNCCNWSNAA 323
   :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :|

QY 61 CTAAGCCAGGTTCTTGTCCTATTATCTTGATTGCTGCTGCTATGTTAAACCCCACTAACC 120
   :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :|
Db 324 RWSNCCNTGGGTNGARWSNGARTAYTNGAYTAYTNTTYGARTNGARCCNGCNCNC 383
   :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :|

QY 121 GTTGTGTTG 128
   :| :| :|
Db 384 NGTNYTNG 391
   :| :| :|

RESULT 67
US-09-995-898A-28
; Sequence 28, Application US/09995898A
; Publication No. US20030027253A1
; GENERAL INFORMATION:
; APPLICANT: Presnell, Scott R.
; APPLICANT: Xu, Wenfeng
; APPLICANT: No. US20030027253A1ak, Julia E.
; APPLICANT: Whitmore, Theodore E.
; APPLICANT: Grant, Francis J.
; TITLE OF INVENTION: CYTOKINE RECEPTOR ZCYTOR19
; FILE REFERENCE: 00-108
; CURRENT APPLICATION NUMBER: US/09/995,898A
; CURRENT FILING DATE: 2001-11-28
; PRIOR APPLICATION NUMBER: US 60/253,561
; PRIOR FILING DATE: 2000-11-28
; PRIOR APPLICATION NUMBER: US 60/267,211
; PRIOR FILING DATE: 2001-02-07
; NUMBER OF SEQ ID NOS: 50
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 28
; LENGTH: 1560
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Degenerate Polynucleotide sequence of SEQ ID
; OTHER INFORMATION: NO:19
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)...(1560)
; OTHER INFORMATION: n = A,T,C or G
US-09-995-898A-28

Query Match      12.4%; Score 25.6; DB 9; Length 1560;
Best Local Similarity 28.9%; Pred. No. 41;
Matches 37; Conservative 24; Mismatches 67; Indels 0; Gaps 0;

QY 1 AATTCGAGCTCGGTACCATACCTGCATATGCTCAAGAACCAAGTTAAAGTCCCTGTCTA 60
   :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :|
Db 264 RAARCARGAYTNTAYAAAYARTTYAARGNGMGTNMGACNGTNSNCCNWSNAA 323
   :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :|

QY 61 CTAAGCCAGGTTCTTGTCCTATTATCTTGATTGCTGCTGCTATGTTAAACCCCACTAACC 120
   :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :|
Db 324 RWSNCCNTGGGTNGARWSNGARTAYTNGAYTAYTNTTYGARTNGARCCNGCNCNC 383
   :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :|

QY 121 GTTGTGTTG 128
   :| :| :|
Db 384 NGTNYTNG 391
   :| :| :|

RESULT 68
```

```

US-09-764-864-359
; Sequence 359, Application US/09764864
; Patent No. US20020132753A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PTZ23
; CURRENT APPLICATION NUMBER: US/09/764,864
; CURRENT FILING DATE: 2001-01-17
; Prior application data removed - consult PALM or file wrapper
; NUMBER OF SEQ ID NOS: 1792
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 359
; LENGTH: 2211
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-764-864-359

Query Match      12.4%; Score 25.6; DB 10; Length 2211;
Best Local Similarity 62.5%; Pred. No. 48;
Matches 40; Conservative 0; Mismatches 24; Indels 0; Gaps 0;

QY 131 GGACACTGATTGTCCAGGTATCAAAAAGTGTGTGAAGGTTCTCTCGGTATGGCTTGT 190
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 2002 GAACGTGGTGGTCCCAAGGTATTTAAAGGGCTTTCTCTGCTCTCTCTCTTGTATT 2061
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

QY 191 CGTT 194
   ||
Db 2062 AATT 2065
   ||

RESULT 69
US-09-764-877-3301
; Sequence 3301, Application US/09764877
; Patent No. US20020147140A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PC005
; CURRENT APPLICATION NUMBER: US/09/764,877
; CURRENT FILING DATE: 2001-01-17
; Prior application data removed - refer to PALM or file wrapper
; NUMBER OF SEQ ID NOS: 4031
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 3301
; LENGTH: 6134
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-764-877-3301

Query Match      12.4%; Score 25.6; DB 10; Length 6134;
Best Local Similarity 49.3%; Pred. No. 79;
Matches 67; Conservative 0; Mismatches 69; Indels 0; Gaps 0;

QY 10 TCGGTACCATACCTGCATATGCTCAAGAACCAAGTTAAAGTCTGTGTCTACTAAGCCAG 69
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 2631 TCTTTACCTACCTAGCCCATTTGGTTAATGCGTTCTTTCTCAGAGTGTAGTACAATT 2690
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

QY 70 GTTCTGTCTATTATCTTGATTGCTGCTGCTATGTTAAACCCCACTAACCCTTGTTCGA 129
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 2691 GTTTTAAGCTTTTTTTTTTTTGTGTAGTATAATCACATCCCTGTTAAGCCTTCTCTGAG 2750
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

QY 130 AGGACACTGATTGTCC 145
   ||| ||| ||| |||
Db 2751 ATCACACCAAAATGTCC 2766
   ||| ||| ||| |||

RESULT 70
US-09-070-927A-138
; Sequence 138, Application US/09070927A
; Patent No. US20020120116A1
; GENERAL INFORMATION:
; APPLICANT: Charles A. Kunsch
```

```

; LENGTH: 326
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-998-598-2252

Query Match      12.3%; Score 25.4; DB 10; Length 326;
Best Local Similarity 64.4%; Pred. No. 23;
Matches 38; Conservative 0; Mismatches 21; Indels 0; Gaps 0;

QY 60 ACTAAGCCAGTCTTGTCCTATTATCTTGATTCGTTGCGCTATGTTAAACCCACCTAA 118
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 191 ACTCTACCTGCACACTGTCTATTATATTCATCTCTTTTGAAATGTCAACCCCAAGTTA 133

RESULT 72
US-09-796-692-7381
; Sequence 7381, Application US/09796692
; Publication No. US20020198362A1
; GENERAL INFORMATION:
; APPLICANT: Gaiger, Alexander
; APPLICANT: Aigate, Paul A.
; APPLICANT: Mannion, Jane
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE DETECTION, DIAGNOSIS AND
; TITLE OF INVENTION: HEMATOLOGICAL MALIGNANCIES
; FILE REFERENCE: 2077.001200
; CURRENT APPLICATION NUMBER: US/09/796,692
; CURRENT FILING DATE: 2001-03-01
; PRIOR APPLICATION NUMBER: 60/186,126
; PRIOR FILING DATE: 2000-03-01
; PRIOR APPLICATION NUMBER: 60/190,479
; PRIOR FILING DATE: 2000-03-17
; PRIOR APPLICATION NUMBER: 60/200,545
; PRIOR FILING DATE: 2000-04-27
; PRIOR APPLICATION NUMBER: 60/200,303
; PRIOR FILING DATE: 2000-04-28
; PRIOR APPLICATION NUMBER: 60/200,779
; PRIOR FILING DATE: 2000-04-28
; PRIOR APPLICATION NUMBER: 60/200,999
; PRIOR FILING DATE: 2000-05-01
; PRIOR APPLICATION NUMBER: 60/202,084
; PRIOR FILING DATE: 2000-05-04
; PRIOR APPLICATION NUMBER: 60/206,201
; PRIOR FILING DATE: 2000-05-22
; PRIOR APPLICATION NUMBER: 60/218,950
; PRIOR FILING DATE: 2000-07-14
; PRIOR APPLICATION NUMBER: 60/222,903
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: 60/223,416
; PRIOR FILING DATE: 2000-08-04
; PRIOR APPLICATION NUMBER: 60/223,378
; PRIOR FILING DATE: 2000-08-07
; NUMBER OF SEQ ID NOS: 9597
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 7381
; LENGTH: 327
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (6)
; OTHER INFORMATION: n=A,T,C or G
; NAME/KEY: unsure
; LOCATION: (221)
; OTHER INFORMATION: n=A,T,C or G
; NAME/KEY: unsure
; LOCATION: (225)
; OTHER INFORMATION: n=A,T,C or G
; NAME/KEY: unsure
; LOCATION: (260)
; OTHER INFORMATION: n=A,T,C or G
; NAME/KEY: unsure
; LOCATION: (320)
; OTHER INFORMATION: n=A,T,C or G

```

```

; PRIOR APPLICATION NUMBER: 60/218,950
; PRIOR FILING DATE: 2000-07-14
; PRIOR APPLICATION NUMBER: 60/222,903
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: 60/223,416
; PRIOR FILING DATE: 2000-08-04
; PRIOR APPLICATION NUMBER: 60/223,378
; PRIOR FILING DATE: 2000-08-07
; NUMBER OF SEQ ID NOS: 9597
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 7381
; LENGTH: 327
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (6)
; OTHER INFORMATION: n=A,T,C or G
; NAME/KEY: unsure
; LOCATION: (221)
; OTHER INFORMATION: n=A,T,C or G
; NAME/KEY: unsure
; LOCATION: (225)
; OTHER INFORMATION: n=A,T,C or G
; NAME/KEY: unsure
; LOCATION: (260)
; OTHER INFORMATION: n=A,T,C or G
; NAME/KEY: unsure
; LOCATION: (320)
; OTHER INFORMATION: n=A,T,C or G
;

```


Sun Feb 16 09:12:56 2003

OTHER INFORMATION: n equals a,t,g, or c
NAME/KEY: misc feature
LOCATION: (450)
OTHER INFORMATION: n equals a,t,g, or c
NAME/KEY: misc feature
LOCATION: (451)
OTHER INFORMATION: n equals a,t,g, or c
NAME/KEY: misc feature
LOCATION: (453)
OTHER INFORMATION: n equals a,t,g, or c
NAME/KEY: misc feature
LOCATION: (454)
OTHER INFORMATION: n equals a,t,g, or c
NAME/KEY: misc feature
LOCATION: (455)
OTHER INFORMATION: n equals a,t,g, or c
NAME/KEY: misc feature
LOCATION: (456)
OTHER INFORMATION: n equals a,t,g, or c
NAME/KEY: misc feature
LOCATION: (462)
OTHER INFORMATION: n equals a,t,g, or c
NAME/KEY: misc feature
LOCATION: (474)
OTHER INFORMATION: n equals a,t,g, or c
NAME/KEY: misc feature
LOCATION: (493)
OTHER INFORMATION: n equals a,t,g, or c
NAME/KEY: misc feature
LOCATION: (494)
OTHER INFORMATION: n equals a,t,g, or c
US-09-783-590-7280

Query Match 12.3%; Score 25.4; DB 10; Length 504;
Best Local Similarity 55.3%; Pred. No. 28;
Matches 47; Conservative 0; Mismatches 38; Indels 0; Gaps 0;
QY 66 CCAGGTTCTTGTCTTATCTTATCTTGTTCGCTATGTTAAACCCACCTAACCGTTGT 125
Db 294 CCAGGTTCTTGTCTTCTCCNCACCTGTGTATGCATAGGTTTANCCGCTTCACCTGTTG 235
QY 126 TTGAAGGACACTGATGTCCAGGTA 150
Db 234 CTGGGGAACACTGCTCTCCGGGGA 210

RESULT 74
US-09-822-830A-205
Sequence 205, Application US/09822830A
Patent No. US20020142952A1
GENERAL INFORMATION:
APPLICANT: Genetics Institute, Inc.
APPLICANT: Wong, Gordon G.
APPLICANT: Clark, Hilary
APPLICANT: Fechtel, Kim
APPLICANT: Agostino, Michael J.
APPLICANT: Howes, Steven H.
APPLICANT: Resnick, Richard J.
APPLICANT: Gulukota, Kamalakara
APPLICANT: Graham, James R.
TITLE OF INVENTION: POLYNUCLEOTIDES ENCODING NOVEL SECRETED PROTEINS
FILE REFERENCE: GIN 6402
CURRENT APPLICATION NUMBER: US/09/822,830A
CURRENT FILING DATE: 2001-03-29
PRIOR APPLICATION NUMBER: 60/195,604
PRIOR FILING DATE: 2000-04-06
NUMBER OF SEQ ID NOS: 631
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 205
LENGTH: 1026
TYPE: DNA
ORGANISM: Homo sapiens
US-09-822-830A-205

Query Match 12.3%; Score 25.4; DB 10; Length 1026;
Best Local Similarity 52.3%; Pred. No. 39;
Matches 56; Conservative 0; Mismatches 51; Indels 0; Gaps 0;
QY 23 TGCATATGCTCAAGAACCAAGGTTAAAGGTCCTGTGTCTACTAAGCCAGGTTCTTGTCTCTAT 82
Db 195 TGCATTTGCTACTTATCTCTGTGCAGGAACCAATCTCTCTCCACCTATCCCTTGTCTCTCC 254
QY 83 TATCTTGATTCGTTGGCTATGTTAAACCCACCTAACCGTTGTTGA 129
Db 255 CCACTTGACTGACTCTCTCTGTCTCTTCAAATCTTAGCTTGAGATGA 301

RESULT 75
US-08-927-939-55
Sequence 55, Application US/08927939
Patent No. US20010006640A1
GENERAL INFORMATION:
APPLICANT: Grainger, David J.
APPLICANT: Tatalick, Lauen Marie
TITLE OF INVENTION: Compounds and methods to inhibit or
TITLE OF INVENTION: augment an inflammatory response.
FILE REFERENCE: 295.022US1
CURRENT APPLICATION NUMBER: US/08/927,939
CURRENT FILING DATE: 1997-09-11
NUMBER OF SEQ ID NOS: 83
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 55
LENGTH: 1041
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: CDS
LOCATION: (18)...(338)
US-08-927-939-55

Query Match 12.3%; Score 25.4; DB 8; Length 1041;
Best Local Similarity 64.4%; Pred. No. 40;
Matches 38; Conservative 0; Mismatches 21; Indels 0; Gaps 0;
QY 60 ACTAAGCCAGGTTCTTGTCTTATCTTATCTTGTTCGCTATGTTAAACCCACCTAA 118
Db 562 ACTCTACCTGCACACTGCTCTATATATATATCTTCTTTTGAATGTCACCCCAAGTTA 620

Search completed: February 16, 2003, 00:33:05
Job time : 376 secs

GenCore version 5.1.4_p5_4578
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 15, 2003, 21:13:02 ; Search time 194 Seconds
(without alignments)
2391.297 Million cell updates/sec

Title: 09-833799-13A
Perfect score: 206
Sequence: 1 AATTCGAGCTCGGTACCATA.....GTTTCGTTCCACAATAATAG 206

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 2185239 seqs, 1125999159 residues
Total number of hits satisfying chosen parameters: 4370478

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 100 summaries

Database :				N_Geneseq_101002:*	
1:	/SIDS2/gcgdata/	geneseq/geneseq-n-emb1/NA1980.DAT:*			
2:	/SIDS2/gcgdata/	geneseq/geneseq-n-emb1/NA1981.DAT:*			
3:	/SIDS2/gcgdata/	geneseq/geneseq-n-emb1/NA1982.DAT:*			
4:	/SIDS2/gcgdata/	geneseq/geneseq-n-emb1/NA1983.DAT:*			
5:	/SIDS2/gcgdata/	geneseq/geneseq-n-emb1/NA1984.DAT:*			
6:	/SIDS2/gcgdata/	geneseq/geneseq-n-emb1/NA1985.DAT:*			
7:	/SIDS2/gcgdata/	geneseq/geneseq-n-emb1/NA1986.DAT:*			
8:	/SIDS2/gcgdata/	geneseq/geneseq-n-emb1/NA1987.DAT:*			
9:	/SIDS2/gcgdata/	geneseq/geneseq-n-emb1/NA1988.DAT:*			
10:	/SIDS2/gcgdata/	geneseq/geneseq-n-emb1/NA1989.DAT:*			
11:	/SIDS2/gcgdata/	geneseq/geneseq-n-emb1/NA1990.DAT:*			
12:	/SIDS2/gcgdata/	geneseq/geneseq-n-emb1/NA1991.DAT:*			
13:	/SIDS2/gcgdata/	geneseq/geneseq-n-emb1/NA1992.DAT:*			
14:	/SIDS2/gcgdata/	geneseq/geneseq-n-emb1/NA1993.DAT:*			
15:	/SIDS2/gcgdata/	geneseq/geneseq-n-emb1/NA1994.DAT:*			
16:	/SIDS2/gcgdata/	geneseq/geneseq-n-emb1/NA1995.DAT:*			
17:	/SIDS2/gcgdata/	geneseq/geneseq-n-emb1/NA1996.DAT:*			
18:	/SIDS2/gcgdata/	geneseq/geneseq-n-emb1/NA1997.DAT:*			
19:	/SIDS2/gcgdata/	geneseq/geneseq-n-emb1/NA1998.DAT:*			
20:	/SIDS2/gcgdata/	geneseq/geneseq-n-emb1/NA1999.DAT:*			
21:	/SIDS2/gcgdata/	geneseq/geneseq-n-emb1/NA2000.DAT:*			
22:	/SIDS2/gcgdata/	geneseq/geneseq-n-emb1/NA2001A.DAT:*			
23:	/SIDS2/gcgdata/	geneseq/geneseq-n-emb1/NA2001B.DAT:*			
24:	/SIDS2/gcgdata/	geneseq/geneseq-n-emb1/NA2002.DAT:*			

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	206	100.0	206	11	AAQ06819
2	127	61.7	8598	19	AAV28852
3	125.4	60.9	177	15	AAQ56674
4	125.4	60.9	177	15	AAQ56676
5	125.2	60.8	480	24	ABK45842
6	124.6	60.5	434	21	AAC68808
7	124.6	60.5	504	11	AAQ06820
8	124.6	60.5	571	24	ABQ58818
9	123.8	60.1	177	15	AAQ56675

10	123.2	59.8	2309	24	ABK83819	Human cDNA differe
11	123.2	59.8	2309	24	ABL65816	Lung cancer relate
12	123.2	59.8	2309	24	ABL66488	Lung cancer relate
13	121	58.7	321	13	AAQ28757	Partial sequence o
14	66.2	32.1	737	15	AAQ44862	SPAI gene. Sus sc
15	45.4	22.0	321	24	ABK88016	DNA encoding human
16	45.4	22.0	1525	24	ABK88022	DNA sequence encod
17	45.4	22.0	1525	24	ABK88025	DNA sequence encod
18	42.6	20.7	321	21	AAC97529	DNA encoding a pro
19	39	18.9	194	10	AAN90354	Sequence encoding
20	38	18.4	656	24	ABQ60721	Human colon cancer
21	34.6	16.8	396	24	ABQ73674	Murine SLPI sense
22	34.6	16.8	396	24	ABQ73677	Murine SLPI antise
23	34.6	16.8	684	20	AAI18516	Mouse IMC carcinom
24	34.6	16.8	691	20	AAI18514	Mouse IMC carcinom
25	34.6	16.8	1114	20	AAI18515	Mouse IMC carcinom
26	33	16.0	416	22	AAI181895	Human polynucleoti
27	32.2	15.6	324	7	AAN60463	Synthetic sequence
28	32.2	15.6	324	7	AAN60464	Synthetic sequence
29	32.2	15.6	324	20	AAI16194	Serine protease in
30	32.2	15.6	324	20	AAI16236	Serine protease in
31	32.2	15.6	324	21	AAC97527	DNA encoding a ser
32	32.2	15.6	324	21	AAC97528	Secretory leukocyt
33	32.2	15.6	324	21	AAC97529	DNA encoding a ser
34	32.2	15.6	324	22	AAI67488	Recombinant serine
35	32.2	15.6	324	22	AAI67489	Recombinant secret
36	32.2	15.6	420	20	AAI16272	DNA sequence of om
37	32.2	15.6	436	21	AAC97624	DNA encoding ompA-
38	32.2	15.6	436	22	AAI67583	DNA sequence codin
39	32.2	15.6	460	20	AAI16271	DNA sequence of om
40	32.2	15.6	460	21	AAC97622	DNA encoding OmpA
41	32.2	15.6	460	22	AAI67582	DNA sequence codin
42	32.2	15.6	546	23	AAS51617	Staphylococcus aur
43	32.2	15.6	549	23	AAS54717	Staphylococcus aur
44	31.6	15.3	16605	22	AAK65489	Human immune/haema
45	31.6	15.3	16605	22	AAK87049	Human immune/haema
46	30.4	14.8	612	23	ABV53540	Human prostate exp
47	30.4	14.8	3496	22	ABA18664	Human nervous syst
48	30.4	14.8	160771	24	ABQ88179	Human osteoblast d
49	30.4	14.8	580073	18	AAT58840	Mycoplasma genital
50	30.2	14.7	450	21	AAC07910	Human secreted pro
51	30.2	14.7	1489	22	AAH64937	Human kinase gene.
52	30.2	14.7	36159	24	ABN85329	Murine replication
53	29.8	14.5	344	22	AAK53717	Blackcurrant fruit
54	29.4	14.3	5150	18	AAT68952	Human prostate exp
55	29.2	14.2	756	24	ABQ89443	Human prostate exp
56	29	14.1	549	22	AAS00193	S. aureus DNA enco
57	29	14.1	549	23	AAS55051	Staphylococcus aur
58	29	14.1	549	23	AAS55319	Staphylococcus aur
59	29	14.1	32145	22	AAK68491	Human immune/haema
60	29	14.1	32145	22	AAK68575	Human immune/haema
61	29	14.1	2365589	24	ABA90521	Genomic sequence o
62	28.8	14.0	1677	21	AAC56088	Eucalyptus grandis
63	28.4	13.8	464	22	ABA43490	Human breast cell
64	28.4	13.8	464	22	ABA53938	Human foetal liver
65	28.4	13.8	464	22	ABA23689	Probe #2155 for ge
66	28.4	13.8	464	22	AAK02208	Human brain expres
67	28.4	13.8	464	22	AAK27655	Human bone marrow
68	28.4	13.8	464	22	AAI12235	Probe #2168 for ge
69	28.4	13.8	464	22	AAI33589	Probe #2275 used t
70	28.4	13.8	464	22	AAI02148	Probe #2139 used t
71	28.4	13.8	464	24	ABS02116	Human genome-deriv
72	28.4	13.8	658	21	AAA16089	Human colon cancer
73	28.4	13.8	1636	24	ABK34741	Human cDNA for nov
74	28.4	13.8	1656	20	AAK60810	Human secreted pro
75	28.4	13.8	1656	22	AAS59216	Human cDNA encodin
76	28.4	13.8	1656	24	ABA90885	Human polynucleoti
77	28.4	13.8	1668	21	AAZ36240	cDNA encoding a bo
78	28.4	13.8	1759	21	AAZ38319	Human transmembran
79	28.4	13.8	1766	19	AAV59748	Human secreted pro
80	28.4	13.8	1915	22	ABA09204	Human neuroendocri
81	28.4	13.8	2454	21	AAC69584	Human secreted pro
82	28.4	13.8	2530	22	AAF93902	Human cDNA encodin

C 83 28.4 13.8 2664 19 AAV59749 Human secreted pro
C 84 28.4 13.8 125910 21 AAC64370 Human KCNQ5 (KCN6q
C 85 28.2 13.7 1473 23 ABL20271 Drosophila melanog
C 86 28.2 13.7 1830 24 AAL45900 A thaliana recepto
C 87 28.2 13.7 4074 23 ABL20270 Drosophila melanog
C 88 28.2 13.7 4488 23 ABL02548 Drosophila melanog
C 89 28 13.6 561 22 AAK91348 Human digestive sy
C 90 28 13.6 1106 22 AAD12603 Human protein havi
C 91 28 13.6 1330 22 AAH34853 Human colon cancer
C 92 28 13.6 1458 22 ABA83379 Human secreted pro
C 93 28 13.6 2054 24 ABQ91970 Human NF-kB activa
C 94 28 13.6 2054 24 ABQ91971 Human NF-kB activa
C 95 28 13.6 2090 22 AAI59408 Human polynucleoti
C 96 28 13.6 2291 22 AAI61194 Human polynucleoti
C 97 28 13.6 1830121 17 AAT42063 Haemophilus influe
C 98 27.8 13.5 338 22 AAK59445 Human immune/haema
C 99 27.8 13.5 626 24 ABQ53920 Oligonucleotide fo
C 100 27.8 13.5 626 24 ABQ53921 Oligonucleotide fo

ALIGNMENTS

RESULT 1
AAQ06819
ID AAQ06819 standard; DNA; 206 BP.
XX
AC AAQ06819;
XX
DT 06-MAR-1991 (first entry)
XX
DE Sequence encoding human leukocyte elastase inhibiting peptide.
XX
KW Emphysema; respiratory distress syndrome; atherosclerosis;
KW arthritis; cystic fibrosis; leukaemia.
XX
OS Homo sapiens.
XX
FH Key
FT CDS
FT Location/Qualifiers
FT 30..203
FT /*tag= a
FT /label= Elastase inhibitor

EP402068-A.
XX
PD 12-DEC-1990.
XX
PF 04-JUN-1990; 90EP-0306037.
XX
PR 02-NOV-1989; 89GB-0024717.
PR 09-JUN-1989; 89GB-0013346.
PR 09-JUN-1989; 89GB-0013349.
PR 25-SEP-1989; 89GB-0021613.
XX
PA (ICIL) IMPERIAL CHEM INDS PLC.
XX
PI Christophers E, Schroder JM, Pioli D, Wiedow O, Edge MD;
XX
DR WPI; 1990-370081/50.
DR P-PSDB; AAR08217.
XX
PT Human leukocyte elastase inhibitor ` is genetically engineered
PT polypeptide for treatment of inflammatory, pulmonary and skin
PT conditions
XX
PS Disclosure; Fig 13; 45pp; English.
XX

CC Gene product has inhibitory activity against human leukocyte
CC elastase (HLE) and may be isolated from psoriatic scales. The
CC sequence may be expressed from a plasmid transformed expression
CC system and may be useful in the prevention of tissue damage
CC associated with emphysema, adult respiratory distress syndrome,
CC psoriasis and bullous dermatoses. Other treatable conditions

CC include atherosclerosis, cystic fibrosis, bronchitis and acute
CC non-lymphoblastic leukaemia.
CC Abs raised to the polypeptide may be used in detection.
XX
SQ Sequence 206 BP; 51 A; 46 C; 44 G; 65 T; 0 other;
Query Match 100.0%; Score 206; DB 11; Length 206;
Best Local Similarity 100.0%; Pred. No. 5.6e-60;
Matches 206; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 AATTCGAGCTCGGTACCATACCTGCATATGCTCAAGAACAGTTAAAGGTCCTGTGTCTA 60
Db 1 AATTCGAGCTCGGTACCATACCTGCATATGCTCAAGAACAGTTAAAGGTCCTGTGTCTA 60
QY 61 CTAAGCCAGGTTCTTGTCTCTATATCTTGTTCGTCGCTATGTTAAACCCACCTAAC 120
Db 61 CTAAGCCAGGTTCTTGTCTCTATATCTTGTTCGTCGCTATGTTAAACCCACCTAAC 120
QY 121 GTTGTTCGAGGACACTGATTGTCAGGTATCAAAAAGTCTGTGAAGGTTCTTCGGGTA 180
Db 121 GTTGTTCGAGGACACTGATTGTCAGGTATCAAAAAGTCTGTGAAGGTTCTTCGGGTA 180
QY 181 TGGCTTGTTCGTTCCACAAATAATAG 206
Db 181 TGGCTTGTTCGTTCCACAAATAATAG 206

RESULT 2
AAV28852
ID AAV28852 standard; cDNA to mRNA; 8598 BP.
XX
AC AAV28852;
XX
DT 04-AUG-1998 (first entry)
XX
DE pPIC9/ELF25L cDNA construct containing the elafin 25L gene.
XX
KW Elafin; elafin 25L gene; Pichia pastoris; expression vector;
KW alcohol oxidase 1; controlling region; promoter; ds.
XX
OS Synthetic.
OS Pichia pastoris.

XX
FH Key
FT CDS
FT Location/Qualifiers
FT 949..1377
FT /*tag= a
FT sig_peptide 949..1203
FT /*tag= b
FT mat_peptide 1204..1374
FT /*tag= c
XX
PN JP10127292-A.
XX
PD 19-MAY-1998.
XX
PF 31-OCT-1996; 96JP-0304233.
XX
PR 31-OCT-1996; 96JP-0304233.
XX
PA (TSUR) TSUMURA & CO.
XX
DR WPI; 1998-340667/30.
DR P-PSDB; AAW57237.
XX
PT New elafin expression vector - used to transform microbial host(s)
PT for production of elafin commercially
XX
PS Example 1; Page 11-15; 21pp; Japanese.

CC The present sequence represents a pPIC9/ELF25L cDNA construct containing
CC the elafin 25L gene, used in an example of the present invention. The
CC present invention describes: (1) an elafin (EL) expression vector
CC comprising a gene encoding EL and a gene encoding alcohol oxidase 1

CC controlling region (AOCR) which controls the expression of EL encoding
CC gene; (2) a microbial host transformed with the vector; (3) DNA fragment
CC for recombinant transformation comprising: (a) a first region homologous
CC to the genomic DNA of the host; (b) a promoter region of (AOCR) gene; (c)
CC a gene encoding signal peptide; (d) a linker designed to secrete a
CC matured elafin; (e) a gene encoding EL; (f) selection marker gene; (g)
CC stop codon, and (h) a second region homologous to a part of the genomic
CC DNA of the host, and (4) a microbial host transformed with the DNA
CC fragment of (3). The microbial hosts can be used to prepare EL
CC commercially. The production of EL by the microbial hosts can be carried
CC on a large scale.
XX
SQ Sequence 8598 BP; 2306 A; 2025 C; 1920 G; 2347 T; 0 other;
Query Match 61.7%; Score 127; DB 19; Length 8598;
Best Local Similarity 82.9%; Pred. No. 1.4e-32;
Matches 145; Conservative 0; Mismatches 30; Indels 0; Gaps 0;
QY 30 GCTCAAGAACCAAGTTAAAGGTCCTGTCTACTAAGCCAGGTTCTTGTCCTATTATCTTG 89
Db 1204 GCTCAAGAACCAAGTTAAGGTCCTGTCTGACCAACCCGGCTCTTGCCCGATTATCCTG 1263
QY 90 ATTCGTTGCGCTATGTTAAACCCACCTAACCGTTGTTTGAAGGACACTGATTGTCAGGT 149
Db 1264 ATCCGCTGGGCTTTGCTGAACCCCGCAACCGTTGTCGAAAGACACTGACTGCCCGGT 1323
QY 150 ATCAAAAAGTGTGTGAAGGTTCTGCGGTATGGCTTGTTCGTTCCACAATAAT 204
Db 1324 ATCAAAAAGTGTGTGAAGGTTCTGCGGTATGGCATGCTTCGTTCCGACAGTAGT 1378
RESULT 3
AAQ56674
ID AAQ56674 standard; DNA; 177 BP.
XX
AC AAQ56674;
XX
DT 23-SEP-1994 (first entry)
XX
DE Elafin derivative Leu25 DNA.
XX
KW Elafin; derivative; mutation; transformation; E. coli; yeast;
KW Bacillus subtilis; elastase inhibitor; oxidation; ss.
XX
OS Synthetic.
XX
PN WO9404697-A.
XX
PD 03-MAR-1994.
XX
PF 11-AUG-1993; 93WO-JP01133.
XX
PR 11-AUG-1992; 92JP-0234085.
XX
PA (TSUR) TSUMURA & CO.
XX
PI Amagaya S, Ishima Y, Kaji A, Okawa N, Yoshida M;
XX
DR WPI; 1994-083211/10.
XX
P-PSDB; AAR48552.
PT Recombinant modified elafin with improved oxidation stability -
PT has pharmaceutical use as an elastase inhibitor
XX
PS Claim 1; Page 23; 35pp; Japanese.
XX
CC The sequences given in AAQ56674-76 encode elafin derivatives. The DNA
CC encoding wild type elafin is mutated, inserted into a suitable
CC vector and then used to transform E. coli, yeast, Bacillus subtilis
CC or animal cells. The modified elafin is expressed when the
CC transformed cells are cultured. The modified elafin are drugs with
CC elastase inhibitor activity. They have improved oxidation stability
CC over natural elafin and thus retain activity better under oxidative
CC

CC conditions.
XX
SQ Sequence 177 BP; 39 A; 49 C; 48 G; 41 T; 0 other;
Query Match 60.9%; Score 125.4; DB 15; Length 177;
Best Local Similarity 82.3%; Pred. No. 1.2e-32;
Matches 144; Conservative 0; Mismatches 31; Indels 0; Gaps 0;
QY 30 GCTCAAGAACCAAGTTAAAGGTCCTGTCTACTAAGCCAGGTTCTTGTCCTATTATCTTG 89
Db 1 GCACAGGAACCAAGTTAAAGGTCCTGTCTGACCAACCCGGCTCTTGCCCGATTATCCTG 60
QY 90 ATTCGTTGCGCTATGTTAAACCCACCTAACCGTTGTTTGAAGGACACTGATTGTCAGGT 149
Db 61 ATCCGCTGGGCTTTGCTGAACCCCGCAACCGTTGTCGAAAGACACTGACTGCCCGGT 120
QY 150 ATCAAAAAGTGTGTGAAGGTTCTGCGGTATGGCTTGTTCGTTCCACAATAAT 204
Db 121 ATCAAAAAGTGTGTGAAGGTTCTGCGGTATGGCATGCTTCGTTCCGACAGTAGT 175
RESULT 4
AAQ56676
ID AAQ56676 standard; DNA; 177 BP.
XX
AC AAQ56676;
XX
DT 23-SEP-1994 (first entry)
XX
DE Elafin derivative Ile25 DNA.
XX
KW Elafin; derivative; mutation; transformation; E. coli; yeast;
KW Bacillus subtilis; elastase inhibitor; oxidation; ss.
XX
OS Synthetic.
XX
PN WO9404697-A.
XX
PD 03-MAR-1994.
XX
PF 11-AUG-1993; 93WO-JP01133.
XX
PR 11-AUG-1992; 92JP-0234085.
XX
PA (TSUR) TSUMURA & CO.
XX
PI Amagaya S, Ishima Y, Kaji A, Okawa N, Yoshida M;
XX
DR WPI; 1994-083211/10.
XX
P-PSDB; AAR48554.
PT Recombinant modified elafin with improved oxidation stability -
PT has pharmaceutical use as an elastase inhibitor
XX
PS Claim 1; Page 25; 35pp; Japanese.
XX
CC The sequences given in AAQ56674-76 encode elafin derivatives. The DNA
CC encoding wild type elafin is mutated, inserted into a suitable
CC vector and then used to transform E. coli, yeast, Bacillus subtilis
CC or animal cells. The modified elafin is expressed when the
CC transformed cells are cultured. The modified elafin are drugs with
CC elastase inhibitor activity. They have improved oxidation stability
CC over natural elafin and thus retain activity better under oxidative
CC conditions.
XX
SQ Sequence 177 BP; 40 A; 49 C; 47 G; 41 T; 0 other;
Query Match 60.9%; Score 125.4; DB 15; Length 177;
Best Local Similarity 82.3%; Pred. No. 1.2e-32;
Matches 144; Conservative 0; Mismatches 31; Indels 0; Gaps 0;
QY 30 GCTCAAGAACCAAGTTAAAGGTCCTGTCTACTAAGCCAGGTTCTTGTCCTATTATCTTG 89

Db 1 GCACAGGAACCAAGTTAAAGTCCGGTGTGCGACCAACACGGGCTCTTGGCCGATTATCCTG 60
Qy 90 ATTCTGTCGCTATGTTAAACCCACCTAACCGTTGTTTGAAGGACACTGATTGTCCAGGT 149
Db 61 ATCCGTCGCTATTTCTGAACCCGCCGAACCGTTGTCTGAAAGACACTGACTGCCCGGGT 120
Qy 150 ATCAAAAAGTGTGTGAAGTTCTCTGCGGTATGGCTTGTTCGTTCCACAATAAT 204
Db 121 ATCAAAAAGTGTGTGAAGTTCTCTGCGGTATGGCTTGTTCGTTCCGCGAGT 175

RESULT 5
ABK45842/c
ID ABK45842 standard; cDNA; 480 BP.
XX
AC ABK45842;
XX
DT 05-JUN-2002 (first entry)
XX
DE cDNA encoding colon tumour protein, SEQ ID No 1393.
XX
KW Human; colon tumour; vaccine; colon cancer; immunogenic;
KW immunotherapy; gene; ss.
XX
OS Homo sapiens.
XX
PN WO200212328-A2.
XX
PD 14-FEB-2002.
XX
PF 31-JUL-2001; 2001WO-US24218.
XX
PR 03-AUG-2000; 2000US-223283P.
PR 28-MAR-2001; 2001US-279763P.
PR 29-JUN-2001; 2001US-302051P.
XX
PA (CORI-) CORIXA CORP.
XX
PI King GE, Meagher MJ, Xu J, Secrist H;
XX WPI; 2002-241739/29.
XX
PT New colon cancer polypeptides and polynucleotides, useful as vaccines,
PT for diagnosing, preventing, and treating colon cancer, and as markers
PT for the progression of cancer -

PS Claim 1; SEQ ID No 1393; 147pp; English.
XX
XX The invention relates to polynucleotides encoding colon tumour proteins.
CC The polynucleotides and encoded polypeptides are useful in pharmaceutical
CC compositions, such as vaccines, for the diagnosis, prevention, and
CC treatment of colon cancer. Polynucleotide sequences may be used as
CC hybridisation probes or primers, and in the design and preparation of
CC ribozyme molecules for inhibiting expression of tumour polypeptides and
CC proteins in tumour cells. The compositions are useful for stimulating an
CC immune response against cancer, particularly for the immunotherapy of
CC colon cancer, and as markers for the progression of cancer.
CC ABK44450-ABK46237 represent coding sequences of human colon tumour
CC proteins of the invention.
CC Note: With the exception of SEQ ID No 1 and 2, the sequence data
CC for this patent did not form part of the printed specification but was
CC supplied by the European Patent Office.
XX
SQ Sequence 480 BP; 113 A; 105 C; 139 G; 120 T; 3 other;

Query Match 60.8%; Score 125.2; DB 24; Length 480;
Best Local Similarity 81.0%; Pred. No. 2e-32;
Matches 145; Conservative 0; Mismatches 34; Indels 0; Gaps 0;
Qy 28 ATGCTCAAGAACCAAGTTAAAGTCCGTGTCTACTAAGCCAGGTTCTTGTCTATATCT 87
Db 371 AAGCGCAAGAGCCAGTCAAAGGTCCAGTNTCCACTAAGCCTGGCTCTGCCCAATTATCT 312

Qy 88 TGATTCGTTGCGCTATGTTAAACCCACCTAACCGTTGTTTGAAGGACACTGATTGTCCAG 147
Db 311 TGATCCGGTGGCCATGTTGAATCCCCCTAACCGTGTCTTGAAGATATACTGACTGCCAG 252
Qy 148 GTATCAAAAAGTGTGTGAAGTTCTCTGCGGTATGGCTTGTTCGTTCCACAATAATAG 206
Db 251 GAATCAAGAAAGTGTGTGAAGGCTCTTGGGGGATGGCCTGTTTCGTTCCCAAGTGAGAG 193

RESULT 6
AAC68808/c
ID AAC68808 standard; cDNA; 434 BP.
XX
AC AAC68808;
XX
DT 21-FEB-2001 (first entry)
XX
DE Human head/neck tumour related protein partial coding sequence #5.
XX
KW Head tumour; neck tumour; lung cancer; vaccine; cancer therapy; ss.
XX
OS Homo sapiens.
XX
PN WO2000065053-A2.
XX
PD 02-NOV-2000.
XX
PF 21-APR-2000; 2000WO-US10687.
XX
PR 23-APR-1999; 99US-0130906.
PR 20-APR-2000; 2000US-0533870.
XX
PA (CORI-) CORIXA CORP.
XX
PI Wang T, Dillon DC;
XX
DR WPI; 2000-687345/67.
XX

PT Novel polypeptides comprising immunogenic portion of head and neck
PT tumour protein useful for treating, diagnosing and monitoring cancer
PT such as head, neck and lung cancer -

PS Claim 3; Page 71; 77pp; English.
XX
XX The present invention relates to a number of nucleic acid sequences which
CC encode proteins associated with head, neck and lung tumours. These
CC tumours are often not diagnosed until they have spread, and of those who
CC survive, most must endure alterations in facial and neck appearance as
CC well as changes in speech, sight, smell, chewing, swallowing and taste
CC perception. The coding sequences given, and the proteins they encode, can
CC be used in the diagnosis, treatment and vaccination against cancer,
CC particularly papillary and follicular carcinomas, papillary tumours,
CC follicular adenoma, parathyroid hyperplasia, parotid cancer, lip cancer,
CC squamous cell cancer of the tongue, oral tongue cancers and larynx
CC cancer.
XX
SQ Sequence 434 BP; 103 A; 103 C; 133 G; 94 T; 1 other;

Query Match 60.5%; Score 124.6; DB 21; Length 434;
Best Local Similarity 81.0%; Pred. No. 3.1e-32;
Matches 145; Conservative 0; Mismatches 34; Indels 0; Gaps 0;
Qy 28 ATGCTCAAGAACCAAGTTAAAGTCCGTGTCTACTAAGCCAGGTTCTTGTCTATATCT 87
Db 320 AAGCGCAAGAGCCAGTCAAAGGTCCAGTCTCCACTAAGCCTGGCTCTGCCCAATTATCT 261
Qy 88 TGATTCGTTGCGCTATGTTAAACCCACCTAACCGTTGTTTGAAGGACACTGATTGTCCAG 147
Db 260 TGATCCGGTGGCCATGTTGAATCCCCCTAACCGTGTCTGTAAGATATACTGACTGCCAG 201
Qy 148 GTATCAAAAAGTGTGTGAAGTTCTCTGCGGTATGGCTTGTTCGTTCCACAATAATAG 206
Db 200 GAATCAAGAAAGTGTGTGAAGGCTCTTGGGGGATGGCCTGTTTCGTTCCCAAGTGAGAG 142

RESULT 7
AAQ06820
ID AAQ06820 standard; cDNA; 504 BP.
XX
AC
AC
AC
XX
DT 06-MAR-1991 (first entry)
XX
DE Sequence encoding human leukocyte elastase inhibiting peptide.
XX
KW HLE; Emphysema; respiratory distress syndrome; atherosclerosis;
KW arthritis; cystic fibrosis; leukaemia; ds.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT CDS 1..157
FT /*tag= b
FT /label= Upstream in-frame coding sequence
FT CDS 121..328
FT /*tag= a
FT /label= Elastase inhibitor
FT intron 144..180
FT /*tag= c
XX
PN EP402068-A.
XX
PD 12-DEC-1990.
XX
PF 04-JUN-1990; 90EP-0306037.
XX
PR 02-NOV-1989; 89GB-0024717.
PR 09-JUN-1989; 89GB-0013346.
PR 09-JUN-1989; 89GB-0013349.
PR 25-SEP-1989; 89GB-0021613.
XX
PA (ICIL) IMPERIAL CHEM INDS PLC.
XX
PI Christophers E, Schroder JM, Pioli D, Wiedow O, Edge MD;
XX
DR WPI; 1990-370081/50.
XX
DR P-PSDB; AAR08217.
XX
PT Human leukocyte elastase inhibitor - is genetically engineered
PT polypeptide for treatment of inflammatory, pulmonary and skin
PT conditions
XX
PS Disclosure; Fig 16; 45pp; English.
XX
CC Gene product has inhibitory activity against human leukocyte
CC elastase (HLE) and may be isolated from psoriatic scales. The
CC sequence may be expressed from a plasmid transformed expression
CC system and may be useful in the prevention of tissue damage
CC associated with emphysema, adult respiratory distress syndrome,
CC psoriasis and bullous dermatoses. Other treatable conditions
CC include atherosclerosis, cystic fibrosis, bronchitis and acute
CC non-lymphoblastic leukaemia.
CC Abs raised to the polypeptide may be used in detection.
XX
SQ Sequence 504 BP; 110 A; 152 C; 120 G; 122 T; 0 other;

Query Match 60.5%; Score 124.6; DB 11; Length 504;
Best Local Similarity 81.0%; Pred. No. 3.3e-32;
Matches 145; Conservative 0; Mismatches 34; Indels 0; Gaps 0;

QY 28 ATGCTCAAGAACCCAGTTAAAGTCTGTGCTACTAAGCCAGGTTCTTGTCTATTATCT 87
DB 156 AAGCGCAAGAGCCAGTCAAGGTCAGTCTCCACTAAGCCTGGCTCTGCCCATATCT 215

QY 88 TGATTCGTTGCGCTATGTTAAACCCACCTAACCGTTGTTTGAAGGACACTGATGTCCAG 147
DB 147

Db 216 TGATCCGGTGGCCCATGTTGAATCCCCCTAACCGCTGCTTGAAAGATACTGACTGCCAG 275
QY 148 GTATCAAAAAGTCTGTGAAGGTTCTCTGCGGTATGGCTTGTTCGTTCCACAATAATAG 206
DB 276 GAATCAAGAAGTCTGTGAAGGCTCTTGCGGGATGGCCTGTTTCGTTCCCCAGTGAGAG 334

RESULT 8
ABQ58818
ID ABQ58818 standard; cDNA; 571 BP.
XX
AC ABQ58818;
XX
DT 02-AUG-2002 (first entry)
XX
DE Human colon cancer related nucleotide sequence SEQ ID NO:2513.
XX
KW Human; colon cancer; cancer; tissue profiling; forensic; mapping;
KW genetic analysis; diagnostic; antisense therapy; gene; ss.
XX
OS Homo sapiens.
XX
PN WO200229086-A2.
XX
PD 11-APR-2002.
XX
PF 02-OCT-2001; 2001WO-US30732.
XX
PR 02-OCT-2000; 2000US-237271P.
XX
PA (FARB) BAYER CORP.
XX
PI Burgess C, Astle JH, Carroll E, Catino TJ, Dwivedi P, Molino GA;
PI Thiagalingam A, Lewis ME;
DR WPI; 2002-426115/45.
XX
PT New isolated nucleic acid that is differentially expressed in cancer
PT tissues useful for determining the presence of colon cancer in a cell
PT or tissue type, and in antisense therapy -
XX
PS Claim 1; Fig 1; 796pp; English.
XX
CC ABQ56306 to ABQ60787 represent isolated nucleic acids (I) differentially
CC expressed in cancer tissues. ABB78993 to ABB79004 represent proteins
CC encoded by the ABQ60776 to ABQ60787 nucleic acid sequences. (I) can be
CC used in antisense therapy. An antibody immunoreactive with a polypeptide
CC encoded by (I) is useful for detecting cancer in a patient sample, and
CC for detecting the presence or absence of a polynucleotide encoded by a
CC nucleic acid which hybridises to (I) in a cell. A probe/primer derived
CC from (I) can be used for determining the presence of a nucleic acid which
CC hybridises to (I), and for determining the phenotype of cells in a sample
CC of cells from a patient. (I) is useful for determining the presence of
CC colon cancer in a cell or tissue type, for determining the presence or
CC state of other type of cancer, in antisense therapy, to generate
CC macroarrays on a solid surface, to identify a chromosome on which the
CC corresponding gene resides, and in tissue profiling, forensics, genetic
CC analysis, mapping and diagnostic applications. (I) can be used to raise
CC antibodies, and to screen for peptide analogues and antagonists.
XX
SQ Sequence 571 BP; 136 A; 155 C; 131 G; 145 T; 4 other;

Query Match 60.5%; Score 124.6; DB 24; Length 571;
Best Local Similarity 81.0%; Pred. No. 3.4e-32;
Matches 145; Conservative 0; Mismatches 34; Indels 0; Gaps 0;

QY 28 ATGCTCAAGAACCCAGTTAAAGTCTGTGCTACTAAGCCAGGTTCTTGTCTATTATCT 87
DB 208 AAGCGCAAGAGCCAGTCAAGGTCAGTCTCCACTAAGCCTGGCTCTGCCCATATCT 267

QY 88 TGATTCGTTGCGCTATGTTAAACCCACCTAACCGTTGTTTGAAGGACACTGATGTCCAG 147
DB 268 TGATCCGGTGGCCCATGTTGAATCCCCCTAACCGCTGCTTGAAAGATACTGACTGCCAG 327


```
CC ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 2309 BP; 595 A; 537 C; 603 G; 572 T; 2 other;

Query Match          59.8%; Score 123.2; DB 24; Length 2309;
Best Local Similarity 81.2%; Pred. No. 1.7e-31;
Matches 143; Conservative 0; Mismatches 33; Indels 0; Gaps 0;

QY 28 ATGCTCAAGAACCCAGTTAAAGGTCCTGTGTCTACTAAGCCAGGTTCTTGTCCTATATCT 87
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1552 AAGCGAAGAGCCAGTCAAAGGTCAGTCTCCACTAAGCTGGCTCCTGCCCAATTATCT 1611

QY 88 TGATTGCTGGCGCTATGTTAAACCCACCTAACCGTTGTTGAAGGACACTGATGTCCAG 147
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1612 TGATCCGGTGGCCATGTTGAATCCCCCTAACCGCTGCTTGAAAGATACTGACTGCCAG 1671

QY 148 GTATCAAAAAGTGTGTAAGGTTCTCCTGCGGTATGCTTGAAGGACACTGATGTCCAG 203
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1672 GAATCAAGAAGTGTGTGAAGGCTCTTGCGGGATGGCCTGTTTCGTTCCCCAGTGA 1727

RESULT 11
ABL65816
ID ABL65816 standard; DNA; 2309 BP.
XX
AC ABL65816;
XX
DT 15-MAY-2002 (first entry)
XX
DE Lung cancer related gene sequence SEQ ID NO:4153.
XX
KW Human; cancer; colon; breast; ovary; oesophagus; kidney; thyroid;
KW stomach; lung; prostate; pancreas; carcinoma; antitumour; cancerous;
KW cytostatic; gene therapy; antineoplastic; Wilm's tumour; adenocarcinoma;
KW gene; ds.
XX
OS Homo sapiens.
XX
PN WO200194629-A2.
XX
PD 13-DEC-2001.
XX
PF 30-MAY-2001; 2001WO-US10838.
XX
PR 05-JUN-2000; 2000US-209473P.
PR 05-JUN-2000; 2000US-209531P.
PR 18-SEP-2000; 2000US-233133P.
PR 18-SEP-2000; 2000US-233617P.
PR 20-SEP-2000; 2000US-234009P.
PR 20-SEP-2000; 2000US-234034P.
PR 20-SEP-2000; 2000US-234052P.
PR 22-SEP-2000; 2000US-234509P.
PR 22-SEP-2000; 2000US-234567P.
PR 25-SEP-2000; 2000US-234923P.
PR 25-SEP-2000; 2000US-234924P.
PR 25-SEP-2000; 2000US-235077P.
PR 25-SEP-2000; 2000US-235082P.
PR 25-SEP-2000; 2000US-235134P.
PR 25-SEP-2000; 2000US-235280P.
PR 26-SEP-2000; 2000US-235637P.
PR 26-SEP-2000; 2000US-235638P.
PR 27-SEP-2000; 2000US-235711P.
PR 27-SEP-2000; 2000US-235720P.
PR 27-SEP-2000; 2000US-235840P.
PR 27-SEP-2000; 2000US-235863P.
PR 28-SEP-2000; 2000US-236028P.
PR 28-SEP-2000; 2000US-236032P.
PR 28-SEP-2000; 2000US-236033P.
PR 28-SEP-2000; 2000US-236034P.
PR 28-SEP-2000; 2000US-236109P.
PR 28-SEP-2000; 2000US-236111P.
PR 29-SEP-2000; 2000US-236842P.
PR 29-SEP-2000; 2000US-236891P.
```

```
PR 02-OCT-2000; 2000US-237172P.
PR 02-OCT-2000; 2000US-237173P.
PR 02-OCT-2000; 2000US-237278P.
PR 02-OCT-2000; 2000US-237294P.
PR 02-OCT-2000; 2000US-237295P.
PR 02-OCT-2000; 2000US-237316P.
PR 03-OCT-2000; 2000US-237425P.
PR 03-OCT-2000; 2000US-237598P.
PR 03-OCT-2000; 2000US-237604P.
PR 03-OCT-2000; 2000US-237606P.
PR 03-OCT-2000; 2000US-237608P.
PR 01-NOV-2000; 2000US-244867P.
PR 01-NOV-2000; 2000US-245084P.
XX
PA (AVAL-) AVALON PHARM.
XX
PI Young PE, Augustus M, Carter KC, Ebner R, Endress G, Horrigan S;
PI Soppet DR, Weaver Z;
XX
DR WPI; 2002-188264/24.
XX
PT Screening for anti-neoplastic agent involves exposing cells to a
PT chemical agent to be tested for anti-neoplastic activity, and
PT determining a change in expression of a gene of a signature gene set
XX
PS Claim 1; SEQ ID 4153; 44pp; English.
XX
CC The present invention describes a method (M1) for screening for an
CC anti-neoplastic agent. The method involves exposing cells to a chemical
CC agent to be tested for anti-neoplastic activity, determining a change in
CC expression of at least one gene (I) of a signature gene set, where (I)
CC comprises a sequence (S) selected from 8447 sequences (given in ABL61664
CC to ABL70110), or is at least 95% identical to (S), where a change in
CC expression is indicative of anti-neoplastic activity. (I) has cytostatic
CC activity and can be used in gene therapy. M1 can be used for screening
CC an anti-neoplastic agent, and can be used for producing a product which
CC is the data collected with respect to the anti-neoplastic agent as a
CC result of M1, and the data is sufficient to convey the chemical
CC structure and/or properties of the agent. M1 can be used in the
CC treatment of cancer such as colon, breast, stomach, lung, thyroid,
CC oesophageal, ovarian, kidney, prostate or pancreatic cancer,
CC adenocarcinoma, carcinoma, clear cell cancer, infiltrating ductal cancer,
CC infiltrating lobular cancer, squamous cell carcinoma, neuroendocrine
CC carcinoma, papillary carcinoma and Wilm's tumour.
XX
SQ Sequence 2309 BP; 595 A; 537 C; 603 G; 572 T; 2 other;

Query Match          59.8%; Score 123.2; DB 24; Length 2309;
Best Local Similarity 81.2%; Pred. No. 1.7e-31;
Matches 143; Conservative 0; Mismatches 33; Indels 0; Gaps 0;

QY 28 ATGCTCAAGAACCCAGTTAAAGGTCCTGTGTCTACTAAGCCAGGTTCTTGTCCTATATCT 87
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1552 AAGCGAAGAGCCAGTCAAAGGTCAGTCTCCACTAAGCTGGCTCCTGCCCAATTATCT 1611

QY 88 TGATTGCTGGCGCTATGTTAAACCCACCTAACCGTTGTTGAAGGACACTGATGTCCAG 147
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1612 TGATCCGGTGGCCATGTTGAATCCCCCTAACCGCTGCTTGAAAGATACTGACTGCCAG 1671

QY 148 GTATCAAAAAGTGTGTAAGGTTCTCCTGCGGTATGCTTGAAGGACACTGATGTCCAG 203
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1672 GAATCAAGAAGTGTGTGAAGGCTCTTGCGGGATGGCCTGTTTCGTTCCCCAGTGA 1727

RESULT 12
ABL66488
ID ABL66488 standard; DNA; 2309 BP.
XX
AC ABL66488;
XX
DT 15-MAY-2002 (first entry)
XX
DE Lung cancer related gene sequence SEQ ID NO:4825.
```



```

streptavidin. The remaining single-stranded phagemid cDNAs are
converted to double-stranded form and used to transform bacterial
host cells. The resulting subtracted cDNA library is differentially
screened with total cDNA from normal and tumor cells. This method
produced some 20 additional cloned cDNAs. Also found by this
method were several genes which, on the basis of the partial DNA
sequences appear to be novel sequences not previously entered
into GENBANK. The portion of the cDNAs so sequenced represents
part of the coding region and/or part of the 3' untranslated region
of each cDNA (see Q28749-58).

Sequence 321 BP; 81 A; 80 C; 84 G; 76 T; 0 other;

Query Match      58.7%; Score 121; DB 13; Length 321;
Best Local Similarity 82.2%; Pred. No. 4.6e-31;
Matches 139; Conservative 0; Mismatches 30; Indels 0; Gaps 0;

QY 28 ATGCTCAAGAACCCAGTTAAAGGTCCTGTGCTCTACTAAGCCAGGTTCTTGTCCTATTATCT 87
Db 142 AAGCGCAAGAGCCAGTCAAAGGTCCTCAGTCTCCACTAAGCCCTGGCTCCTGCCCATTAATCT 201

QY 88 TGATTCTGTCGGCTATGTTAAACCCACCTAACCCTGTTGTTTGAGGACACTGATTGTCCAG 147
Db 202 TGATCCGGTGCGCCATGTTGAATCCTCCTAACCCTGCTTGAAAGATACTGACTGCCCCAG 261

QY 148 GTATCAAAAAGTGCTGTGAAGGTTCTCTCGGATGCGCTTGTTGTTTCC 196
Db 262 GAATCAAGAAGTGCTGTGAAGGCTCTTGCGGGATGGCCTGTTGTTTCC 310

RESULT 14
AAQ44862
ID AAQ44862 standard; cDNA to mRNA; 737 BP.
XX
AC AAQ44862;
XX
DT 17-OCT-1994 (first entry)
XX
DE SPAI gene.
XX
KW SPAI; sodium ion, potassium ion, ATP-ase inhibitor peptide; cardiac;
KW vascular disease; PCR; polymerase chain reaction; ss.
XX
OS Sus scrofa.
XX
FH Key Location/Qualifiers
FT CDS 1..570
FT /*tag= a
FT /product= SPAI
XX
PN JP06049098-A.
XX
PD 22-FEB-1994.
XX
PF 29-JUL-1992; 92JP-0202286.
XX
PR 29-JUL-1992; 92JP-0202286.
XX
PA (EISA ) EISAI CO LTD.
XX
DR WPI; 1994-097819/12.
DR P-PSDB; AAR50334.
XX
PT Sodium ion, potassium ion ATP-ase inhibitor peptide and corresp.
PT DNA - useful to treat cardiac and vascular disease
XX
PS Claim 1; Page 6-7; 8pp; Japanese.
XX
CC The sequence (AAQ44862) encodes the SPAI peptide which is useful
CC as a drug for treating cardiac and vascular diseases. The peptide
CC is found as a pre-sequence (AAR50335) which is amplified using
CC primers (AAQ44863-67)
XX

```

CC inhibitor of protease activity. The fusion protein of the invention
CC is useful for inhibiting protease activity associated with a disorder
CC such as emphysema, asthma, chronic obstructive pulmonary disease,
CC cystic fibrosis, otitis media, otitis external or HIV infection, or
CC for treating an individual suffering from or at risk for a disease or
CC disorder involving unwanted protease activity. The proteins are useful
CC for treating dermatological diseases such as atopic dermatitis, eczema
CC and psoriasis, in inflammatory responses to viral infection, and for
CC treating herpes infection, corneal or epidermal ulceration, chronic
CC non-healing wounds, sepsis, rheumatoid arthritis, periodontal disease,
CC tumour metastasis and tumour angiogenesis, gastric ulceration,
CC osteoporosis, Paget's disease, glomerulonephritis, scleroderma, malaria,
CC bacterial infection, Alzheimer's disease, hypertension and muscular
CC dystrophy. The present sequence represents the DNA encoding the human
CC secretory leukoprotease inhibitor used to create the fusion protein
CC of the invention.
XX Sequence 321 BP; 92 A; 67 C; 80 G; 82 T; 0 other;
SQ
Query Match 22.0%; Score 45.4; DB 24; Length 321;
Best Local Similarity 58.5%; Pred. No. 2.1e-05;
Matches 79; Conservative 0; Mismatches 56; Indels 0; Gaps 0;
QY 63 AAGCCAGGTTCTTCCTATTATCTTGATTCGTTGCGCTATGTTAAACCCACCTAACCGT 122
DB 178 AAGCCAGGTAAGTGTCCAGTTACTTACGGTCAATGTTTGATGTTGAACCCACCAACTTC 237
QY 123 TGTGAAAGGACACTGATTGTCAGGTATCAAAAAGTCTGTGAAGGTTCTCGCGGTATG 182
DB 238 TGTGAAATGGACGGTCAATGTAAGAGAGACTTGAAGTGTGTTGGGTATGTGTGTTAAG 297
QY 183 GCTTGTTCGTTCCA 197
DB 298 TCCTGTGTTTCCCCA 312
RESULT 16
ABK88022
ID ABK88022 standard; DNA; 1525 BP.
XX
AC ABK88022;
XX
DT 07-OCT-2002 (first entry)
XX
DE DNA sequence encoding SLAP1 fusion protein.
XX Yeast; alpha factor; gene; ds; Alzheimer's disease; SLAP1;
KW malaria; emphysema; asthma; chronic obstructive pulmonary disease;
KW cystic fibrosis; otitis media; otitis external; HIV; psoriasis; eczema;
KW human immunodeficiency virus; atopic dermatitis; muscular dystrophy;
KW herpes; ulceration; sepsis; rheumatoid arthritis; periodontal disease;
KW tumour metastasis; tumour angiogenesis; osteoporosis; Paget's disease;
KW glomerulonephritis; scleroderma; hypertension.
XX
OS Homo sapiens.
OS Synthetic.
XX
FH Key Location/Qualifiers
FT RBS 6..8
FT /*tag= a
FT /*standard_name= "Ribosome binding site"
FT CDS 9..1520
FT /*tag= b
FT /*product= "SLAP1 fusion protein"
FT misc_feature 12..332
FT /*tag= c
FT /*note= "SLPI coding region"
FT misc_feature 333-335
FT /*tag= d
FT /*note= "linking codon"
FT misc_feature 336..1517
FT /*tag= e
FT /*note= "AAT coding region"

XX WO200250287-A2.
PN
XX
XX 27-JUN-2002.
PD
XX
PF 18-DEC-2001; 2001WO-US49256.
XX
XX 18-DEC-2000; 2000US-256699P.
PR 20-NOV-2001; 2001US-331966P.
XX
PA (ARRI-) ARRIVA PHARM INC.
XX
XX Barr PJ, Gibson HL, Pemberton P;
PI
XX WPI; 2002-500631/53.
DR P-PSDB; AAU99881.
DR
XX
PT Novel fusion protein useful for inhibiting protease activity associated
PT with a disorder such as emphysema, asthma, comprises a first protease
PT inhibitor comprising alpha 1-antitrypsin and a second protease
PT inhibitor -
XX
PS Example 1; Page 73-73; 134pp; English.
XX
CC This invention relates to a novel fusion protein comprising a first
CC protease inhibitor comprising an alpha 1-antitrypsin or its functionally
CC active portion and a second protease inhibitor or its functionally
CC active protein. The fusion proteins of the invention may act as an
CC inhibitor of protease activity. The fusion protein of the invention
CC is useful for inhibiting protease activity associated with a disorder
CC such as emphysema, asthma, chronic obstructive pulmonary disease,
CC cystic fibrosis, otitis media, otitis external or HIV infection, or
CC for treating an individual suffering from or at risk for a disease or
CC disorder involving unwanted protease activity. The proteins are useful
CC for treating dermatological diseases such as atopic dermatitis, eczema
CC and psoriasis, in inflammatory responses to viral infection, and for
CC treating herpes infection, corneal or epidermal ulceration, chronic
CC non-healing wounds, sepsis, rheumatoid arthritis, periodontal disease,
CC tumour metastasis and tumour angiogenesis, gastric ulceration,
CC osteoporosis, Paget's disease, glomerulonephritis, scleroderma, malaria,
CC bacterial infection, Alzheimer's disease, hypertension and muscular
CC dystrophy. The present sequence represents the DNA encoding the
CC SLAP1 fusion protein of the invention.
XX
SQ Sequence 1525 BP; 467 A; 286 C; 314 G; 458 T; 0 other;
Query Match 22.0%; Score 45.4; DB 24; Length 1525;
Best Local Similarity 58.5%; Pred. No. 3.7e-05;
Matches 79; Conservative 0; Mismatches 56; Indels 0; Gaps 0;
QY 63 AAGCCAGGTTCTTCCTATTATCTTGATTCGTTGCGCTATGTTAAACCCACCTAACCGT 122
DB 189 AAGCCAGGTAAGTGTCCAGTTACTTACGGTCAATGTTTGATGTTGAACCCACCAACTTC 248
QY 123 TGTGAAAGGACACTGATTCCTCCAGGTATCAAAAAGTCTGTGAAGGTTCTTCGCGGTATG 182
DB 249 TGTGAAATGGACGGTCAATGTAAGAGAGACTTGAAGTGTGTTGGGTATGTGTGTTAAG 308
QY 183 GCTTGTTCGTTCCA 197
DB 309 TCCTGTGTTTCCCCA 323
RESULT 17
ABK88025
ID ABK88025 standard; DNA; 1525 BP.
XX
AC ABK88025;
XX
DT 07-OCT-2002 (first entry)
XX
DE DNA sequence encoding rSLAP1 fusion protein.
XX

CC destruction, e.g. emphysema, arthritis, glomerulonephritis,
CC periodontitis, muscular dystrophy or tumour invasion. It is also useful
CC for modulating protease activity. The DNAs are useful for producing the
CC proteins, especially by recombinant methods.
XX
SQ Sequence 321 BP; 60 A; 54 C; 60 G; 44 T; 103 other;
Query Match 20.7%; Score 42.6; DB 21; Length 321;
Best Local Similarity 42.5%; Pred. No. 0.00019;
Matches 57; Conservative 16; Mismatches 61; Indels 0; Gaps 0;
QY 63 AAGCCAGGTTCTTGTCCCTATTATCTTGGTTCGCTATGTTAAACCCACCTAACCGT 122
||:||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| :
Db 178 AARCCNGGNAARTGCCNGTNACNTAYGGNCARTGYCTNATGCTNAAAYCCNCCNAAAYTTY 237
QY 123 TGTGTTGAAGGACACTGATGTCCTCAGGTATCAAAAAGTCTGTGAAGTTCTCTGCGGTATG 182
||:||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| :
Db 238 TNGARATGGAYGGNCARTGYAARCGNGAYCTNAAARTGYTGATGGGNATNTGYGGNAAR 297
QY 183 GCTTGTGTTTCGTTCC 196
||:||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| :
Db 298 TCNTGYGTNTCNC 311
RESULT 19
AAN90354
ID AAN90354 standard; DNA; 194 BP.
XX
AC AAN90354;
XX
DT 01-NOV-1989 (first entry)
XX
DE Sequence encoding fragment of human polymorphonuclear leukocyte
DE elastase inhibiting protein.
XX
KW Elastase inhibitor; human polymorphonuclear leukocyte elastase inhibiting
KW protein.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT misc_feature 1..5 /tag= a
FT misc_feature 7..12 /tag= b
FT misc_feature 25..30 /tag= c
FT misc_feature 56..61 /tag= d
FT misc_feature 120..125 /tag= e
FT misc_feature 185..189 /tag= f
FT misc_feature 190..194 /tag= g
XX
PN WO8906239-A.
XX
PD 13-JUL-1989.
XX
PF 28-DEC-1988; 88WO-JP01342.
XX
PR 28-DEC-1987; 87JP-0330219.
XX
PA (TEIJ) TEIJIN LTD.
XX
PI Sugiyama T, Kamimura T, Masuda K, Okada M, Ohtsuka E;
XX
DR WPI; 1989-220549/30.
XX
P-PSDB; AAP90126, AAP90384.
PT Elastase-inhibiting peptide of low serine protease-inhibiting activity
PT - contains C-terminal portion of human polymorphonuclear leukocyte

PT elastase inhibiting protein.
XX
PS Disclosure; fig 2; 55pp; Japanese.
XX
CC The sequence encodes a fragment of human polymorphonuclear leukocyte
CC elastase inhibiting protein (Asn 55-Ala 107). See AAP90126 for encoded
CC peptide. Misc. features a-g are BamHI, Sali, MluI, NdeI, BglII, XhoI and
CC PstI sites resp. See also AAP90384.
XX
SQ Sequence 194 BP; 48 A; 42 C; 53 G; 51 T; 0 other;
Query Match 18.9%; Score 39; DB 10; Length 194;
Best Local Similarity 55.6%; Pred. No. 0.0026;
Matches 75; Conservative 0; Mismatches 60; Indels 0; Gaps 0;
QY 62 TAAGCCAGGTTCTTGTCCCTATTATCTTGGTTCGCTATGTTAAACCCACCTAACCG 121
||:||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| :
Db 33 TAAACCGGGTAAATGTCCGGTTACATATGGTCAGTGTCTGATGCTGACCCGCCGAACTT 92
QY 122 TTGTTTGAAGGACACTGATGTCCTCAGGTATCAAAAAGTCTGTGAAGTTCTCTGCGGTAT 181
||:||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| :
Db 93 CTGTGAAATGGACGGTCAGTGTAAACGAGATCTGAAATGTTGTATGGGTATGTGTGTAA 152
QY 182 GGCTTGTTCGTTCC 196
||:||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| :
Db 153 ATCTGTGTTTCTCC 167
RESULT 20
ABQ60721/C
ID ABQ60721 standard; cDNA; 656 BP.
XX
AC ABQ60721;
XX
DT 02-AUG-2002 (first entry)
XX
DE Human colon cancer related nucleotide sequence SEQ ID NO:4416.
XX
KW Human; colon cancer; cancer; tissue profiling; forensic; mapping;
KW genetic analysis; diagnostic; antisense therapy; gene; ss.
XX
OS Homo sapiens.
XX
PN WO200229086-A2.
XX
PD 11-APR-2002.
XX
PF 02-OCT-2001; 2001WO-US30732.
XX
PR 02-OCT-2000; 2000US-237271P.
XX
PA (FARB) BAYER CORP.
XX
PI Burgess C, Astle JH, Carroll E, Catino TJ, Dwivedi P, Molino GA;
PI Thiagalingam A, Lewis ME;
XX
DR WPI; 2002-426115/45.
XX
PT New isolated nucleic acid that is differentially expressed in cancer
PT tissues useful for determining the presence of colon cancer in a cell
PT or tissue type, and in antisense therapy -
XX
PS Claim 1; Fig 1; 796pp; English.
XX
CC ABQ56306 to ABQ60787 represent isolated nucleic acids (I) differentially
CC expressed in cancer tissues. ABB78993 to ABB79004 represent proteins
CC encoded by the ABQ60776 to ABQ60787 nucleic acid sequences. (I) can be
CC used in antisense therapy. An antibody immunoreactive with a polypeptide
CC encoded by (I) is useful for detecting cancer in a patient sample, and
CC for detecting the presence or absence of a polynucleotide encoded by a
CC nucleic acid which hybridises to (I) in a cell. A probe/primer derived
CC from (I) can be used for determining the presence of a nucleic acid which
CC hybridises to (I), and for determining the phenotype of cells in a sample

CC of cells from a patient. (I) is useful for determining the presence of
CC colon cancer in a cell or tissue type, for determining the presence or
CC state of other type of cancer, in antisense therapy, to generate
CC macroarrays on a solid surface, to identify a chromosome on which the
CC corresponding gene resides, and in tissue profiling, forensics, genetic
CC analysis, mapping and diagnostic applications. (I) can be used to raise
CC antibodies, and to screen for peptide analogues and antagonists.
XX
SQ Sequence 656 BP; 96 A; 91 C; 107 G; 256 T; 106 other;

Query Match 18.4%; Score 38; DB 24; Length 656;
Best Local Similarity 65.7%; Pred. No. 0.0088;
Matches 44; Conservative 0; Mismatches 23; Indels 0; Gaps 0;

QY 139 ATTGTCAGGTATCAAAAAGTGTGTGAAGGTTCTCTGCGGATGGCTTGTTCGTTCCAC 198
DB 285 ACTGNNCAGGAATCAAGAANTNNTGCAAGGNTCTAGNGGATGCCGATTTGTTCCCC 226

QY 199 AATAATA 205
DB 225 ANNGANA 219

RESULT 21
ABQ73674
ID ABQ73674 standard; DNA; 396 BP.
XX
AC ABQ73674;
XX
DT 04-OCT-2002 (first entry)
XX
DE Murine SLPI sense riboprobe SEQ ID NO:20.
XX
KW Murine; serum amyloid; SAA3; riboprobe; LST-1; follistatin-like 1;
KW FST1; SLPI; secretory leukocyte protease inhibitor; legumain; SOCS3;
KW rheumatoid arthritis; CISH3; RAGE; AGER; LY117; hydroxymethyl glutarate;
KW HMG-1; glucocorticoid leucine zipper; GILZ; PTPN-18; GADD-45A; GADD-45B;
KW PRSC1; lipocalin 2; Lcn2; glucose phosphate isomerase; GPI; SpIL;
KW serine protease inhibitor; TSG-6; ss.
XX
OS Mus sp.
XX
PN WO200248310-A2.
XX
PD 20-JUN-2002.
XX
PF 17-DEC-2001; 2001WO-US48968.
XX
PR 15-DEC-2000; 2000US-255861P.
XX
PA (GENE-) GENETICS INST LLC.
XX
PI Pittman DD, Feldman JL, Shields KM, Trepicchio WL;
XX
DR WPI; 2002-583494/62.
XX
PT Determining difference between expression levels of genes
PT characteristic of rheumatoid arthritis in cell and reference level, by
PT comparing expression levels of the genes determined in a cell with
PT reference level -
XX
PS Example 3; Page 88; 296pp; English.
XX
CC The present invention describes a method (M1) for determining the
CC difference between levels of expression of a number of genes
CC characteristic of rheumatoid arthritis (RA) in cell and reference levels
CC of expression of the genes, comprising: (a) providing RNA from a cell;
CC (b) determining levels of RNA of a number of genes (G) characteristic
CC of RA including a number of genes selected from SOCS3 (CISH3), RAGE
CC (AGER), LST-1 (LY117), serum amyloid (SAA) 1-3, hydroxymethyl glutarate
CC (HMG)-1, S100 A8, A9, and A12, secretory leukocyte protease inhibitor
CC (SLPI), glucocorticoid leucine zipper (GILZ), PTPN-18, GADD-45A and B,
CC Legumain (PRSC1), follistatin-like 1 (FST1), lipocalin 2 (Lcn2), glucose

CC phosphate isomerase (GPI), serine protease inhibitor (SpIL), and TSG-6
CC to obtain the levels of expression of the genes in the cell; and
CC (c) comparing the levels of expression of the genes in the cell to a
CC set of reference levels of expression of the genes, to determine the
CC difference between levels of expression of the genes, to determine the
CC characteristic of RA in the cell and reference levels of expression of
CC the genes. M1 is useful for determining whether a subject has or is
CC likely to develop RA, or for determining whether a therapy for RA is
CC effective in a subject having RA who is receiving the therapy. M1 is
CC also useful for determining the stage of RA in a subject, for
CC determining the efficacy of a therapy in a subject having RA, and for
CC determining the likelihood of success of a particular therapy in subject
CC having RA. The present sequence represents a murine SLPI sense riboprobe,
CC which is used in an example from the present invention.
XX

SQ Sequence 396 BP; 94 A; 98 C; 118 G; 86 T; 0 other;

Query Match 16.8%; Score 34.6; DB 24; Length 396;
Best Local Similarity 53.3%; Pred. No. 0.1;
Matches 73; Conservative 0; Mismatches 64; Indels 0; Gaps 0

QY 51 CCTGTGCTACTAAGCCAGGTTCTTCTCTCTATTATCTTGATTCGCTATGTTAAAC 110
DB 244 CCAGTGTGAGGAAGCCTGGGAGTGGTCAAACTCAGGCAAGATGTATGATGCTTAAC 303
QY 111 CCACCTAACCGTTGTTTGAAGGACACTGATTGTCCAGGTATCAAAAAGTCTGTGAAGGT 170
DB 304 CCTCCCAATGTCTGCCAGAGGACGGCAGTGTACGGCAATAACAAGTCTGTGAGGGT 363
QY 171 TCCTGCGGTATGGCTTG 187
DB 364 ATATGTGGGAAAGTCTG 380

RESULT 22
ABQ73677/c
ID ABQ73677 standard; DNA; 396 BP.

XX
AC ABQ73677;
XX
DT 04-OCT-2002 (first entry)
XX
DE Murine SLPI antisense riboprobe SEQ ID NO:23.

XX Murine; serum amyloid; SAA3; riboprobe; LST-1; follistatin-like 1;
KW FST1; SLPI; secretory leukocyte protease inhibitor; legumain; SOCS3;
KW rheumatoid arthritis; CISH3; RAGE; AGER; LY117; hydroxymethyl glutarate;
KW HMG-1; glucocorticoid leucine zipper; GILZ; PTPN-18; GADD-45A; GADD-45B;
KW PRSC1; lipocalin 2; Lcn2; glucose phosphate isomerase; GPI; SpIL;
KW serine protease inhibitor; TSG-6; ss.
XX
OS Mus sp.

XX
PN WO200248310-A2.
XX
PD 20-JUN-2002.
XX
PF 17-DEC-2001; 2001WO-US48968.
XX
PR 15-DEC-2000; 2000US-255861P.
XX
PA (GENE-) GENETICS INST LLC.

XX Pittman DD, Feldman JL, Shields KM, Trepicchio WL;
XX
DR WPI; 2002-583494/62.
XX
PT Determining difference between expression levels of genes
PT characteristic of rheumatoid arthritis in cell and reference level, by
PT comparing expression levels of the genes determined in a cell with
PT reference level -
XX
PS Example 3; Page 88; 296pp; English.

PT having leukocyte elastase and trypsin inhibitory sites
XX
XX
PS Disclosure; Page 14; 59pp; English.
XX
CC The sequence directs synthesis of a single chain polypeptide serine
CC protease-inhibitor, which believed to have at least 2 active sites,
CC 1 exhibiting leukocyte-elastase-inhibiting properties and the other
CC exhibiting activity against trypsin. See also AAN60464-69 and AAP60562-
CC 66.
XX
SQ Sequence 324 BP; 87 A; 85 C; 88 G; 64 T; 0 other;

Query Match	15.6%;	Score 32.2;	DB 7;	Length 324;
Best Local Similarity	51.8%;	Pred. No. 0.63;		
Matches 73;	Conservative	0;	Mismatches 68;	Indels

Qy	63	AAGCCAGGTTCTTGTCCTATTATCTTGATTCGTTGGCGTATGTTAAACCCACCTAACCGT	1222
Db	178	AAACCCGGGTAAATGCCCGGTAAACCTATGGCCAGTGTCTCATGCTGAACCCCCCGCACTTC	2379

QY 123 TGTTTGAAGGACACTGATTGTCCAGGTATCAAAAAGTGTCTGAAGGTTCTTCTGCGGTATG 182

db 238 TGCGAATGACCGGCCAGTGTAAACAGATCTGAAATGCTGTATGCGGTATGTGCGGCCAAA 297

Qy 183 GCTTGTTCGTTCCACAATAA 203
||| ||| ||| |||
Db 298 TCTTGTGTTTCCCCGGTAAA 318

RESULT 28

ID AAN60464 standard; DNA; 324 BP.

AAN60464;

DT 01-JAN-1980 (first entry)

Synthetic sequence capable of directing microbial synthesis of secretory leukocyte protease-inhibitor.

Serine protease-inhibitor; parotid; leukocyte-elastase-inhibitor;
trypsin-; inhibitor; ss.

Synthetic.

W08603519-A.

19-JUN-1986.

04-DEC-1985: 85WO-US02385

04-DEC-1985: 85WO-US02385

02-DEC-1985: 85US-0803471

(SYNE-) SYNERGEN BIOLOG INC

Bandvonadh PK, Eisenberg SP, Stetler GL, Thompson RC;

WPI: 1986-169458/26.

2025-10-27

New synthetic DNA sequences for directing microbial synthesis - for prodn. of single poly:peptide chain serine protease inhibitor having leukocyte elastase and trypsin inhibitory sites

Disclosure: Page 15: 59pp; English.

The sequence directs synthesis of a secretory leukocyte protease-inhibitor. See also AAN60463, AAN60465-69 and AAP60562 AAP60564-66

Sequence 324 BP: 86 A; 85 C; 87 G; 66 T; 0 other;

XX New serine protease inhibitors and DNA sequences for treating a
PT protease-mediated condition or tissue destruction e.g. emphysema or
PT tumor invasion and for recombinant production of inhibitors -
XX
PS Disclosure; Column 5; 47pp; English.
XX
CC This invention relates to new purified and isolated mammalian serine
CC protease inhibitor proteins which comprise at least 8 cysteine residues
CC and no more than 107 amino acids. The protease inhibitors are capable of
CC inhibiting chymotrypsin and elastase. Sequences AAC97526 - AAC97534,
CC AAC97574 - AAC97581 and AAC97614 - AAC97641 represent DNA encoding the
CC protease inhibitors of the invention, and include oligonucleotide
CC sequences used in the isolation and characterisation of the proteins.
CC Primers AAC97535 - AAC97573 and AAC95582 - AAC97613 are used in the
CC construction of DNA encoding the protease inhibitors. Peptide sequences
CC AAB53098 - AAB53122 represent the protease inhibitor of the invention and
CC various other peptides used in the isolation of the protease inhibitors.
CC The protease inhibitors have cytostatic and anti-inflammatory activity.
CC The serine protease inhibitor protein is useful for treating a
CC protease-mediated condition, which includes protease mediated tissue
CC destruction, e.g. emphysema, arthritis, glomerulonephritis,
CC periodontitis, muscular dystrophy or tumour invasion. It is also useful
CC for modulating protease activity. The DNAs are useful for producing the
CC proteins, especially by recombinant methods.
XX
SQ Sequence 324 BP; 86 A; 85 C; 87 G; 66 T; 0 other;

Query Match 15.6%; Score 32.2; DB 21; Length 324;
Best Local Similarity 51.8%; Pred. No. 0.63;
Matches 73; Conservative 0; Mismatches 68; Indels 0; Gaps 0;

QY 63 AAGCAGGTTCTTGTCCTATTATCTTGATTGCGCTATGTTAAACCCACCTAACCGT 122
Db 178 AAACCGGGTAAATGCCCGGTAACTTATGGCCAGTGTCTGATGCTGAACCCCGCACTTC 237

QY 123 TGTGTTGAAGGACACTGATTGTCAGGTATCAAAAAGTCTGTGAAGTTCTTCGCGGTATG 182
Db 238 TGGGAAATGGACGGCCAGTGTAAACGAGATCTGAAATGCTGTATGGGTATGTGCGGCAA 297

QY 183 GCTTGTTCGTTCCACATAA 203
Db 298 TCTGTGTTTCCCGGTAAAA 318

RESULT 33
AAC97579
ID AAC97579 standard; DNA; 324 BP.
XX
AC AAC97579;
XX
DT 27-FEB-2001 (first entry)
XX
DE DNA encoding a serine protease inhibitory peptide.
XX
KW Serine protease inhibitor; cytostatic; anti-inflammatory; arthritis;
KW protease mediated tissue destruction; emphysema; glomerulonephritis;
KW periodontitis; muscular dystrophy; tumour invasion; chymotrypsin;
KW elastase; ds.
XX
OS Synthetic.
XX
PN US6132990-A.
XX
PD 17-OCT-2000.
XX
PF 07-JUN-1991; 91US-0712354.
XX
PR 03-JAN-1989; 89US-0293042.
PR 06-DEC-1984; 84US-0678822.
PR 02-DEC-1985; 85US-0803471.
PR 29-JUL-1986; 86US-0890526.
PR 30-MAR-1987; 87US-0031846.

PR 04-AUG-1987; 87US-0082962.
XX
PA (AMGE-) AMGEN BOULDER INC.
XX
PI Bandyopadhyay PK, Eisenberg SP, Stetler GL, Thompson RC;
XX
DR WPI; 2000-678667/66.
XX
PT New serine protease inhibitors and DNA sequences for treating a
PT protease-mediated condition or tissue destruction e.g. emphysema or
PT tumor invasion and for recombinant production of inhibitors -
XX
PS Example 2; Column 30-32; 47pp; English.
XX
CC This invention relates to new purified and isolated mammalian serine
CC protease inhibitor proteins which comprise at least 8 cysteine residues
CC and no more than 107 amino acids. The protease inhibitors are capable of
CC inhibiting chymotrypsin and elastase. Sequences AAC97526 - AAC97534,
CC AAC97574 - AAC97581 and AAC97614 - AAC97641 represent DNA encoding the
CC protease inhibitors of the invention, and include oligonucleotide
CC sequences used in the isolation and characterisation of the proteins.
CC Primers AAC97535 - AAC97573 and AAC95582 - AAC97613 are used in the
CC construction of DNA encoding the protease inhibitors. Peptide sequences
CC AAB53098 - AAB53122 represent the protease inhibitor of the invention and
CC various other peptides used in the isolation of the protease inhibitors.
CC The protease inhibitors have cytostatic and anti-inflammatory activity.
CC The serine protease inhibitor protein is useful for treating a
CC protease-mediated condition, which includes protease mediated tissue
CC destruction, e.g. emphysema, arthritis, glomerulonephritis,
CC periodontitis, muscular dystrophy or tumour invasion. It is also useful
CC for modulating protease activity. The DNAs are useful for producing the
CC proteins, especially by recombinant methods.
XX
SQ Sequence 324 BP; 86 A; 85 C; 87 G; 66 T; 0 other;

Query Match 15.6%; Score 32.2; DB 21; Length 324;
Best Local Similarity 51.8%; Pred. No. 0.63;
Matches 73; Conservative 0; Mismatches 68; Indels 0; Gaps 0;

QY 63 AAGCAGGTTCTTGTCCTATTATCTTGATTGCGCTATGTTAAACCCACCTAACCGT 122
Db 178 AAACCGGGTAAATGCCCGGTAACTTATGGCCAGTGTCTGATGCTGAACCCCGCACTTC 237

QY 123 TGTGTTGAAGGACACTGATTGTCAGGTATCAAAAAGTCTGTGAAGTTCTTCGCGGTATG 182
Db 238 TGGGAAATGGACGGCCAGTGTAAACGAGATCTGAAATGCTGTATGGGTATGTGCGGCAA 297

QY 183 GCTTGTTCGTTCCACATAA 203
Db 298 TCTGTGTTTCCCGGTAAAA 318

RESULT 34
AAI67488
ID AAI67488 standard; DNA; 324 BP.
XX
AC AAI67488;
XX
DT 11-FEB-2002 (first entry)
XX
DE Recombinant serine protease inhibitor DNA.
XX
KW Serine protease inhibitor protein; recombinant; leukocyte elastase;
KW trypsin; ds.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT CDS 1..324
FT /tag= a
FT /product= "serine protease inhibitor"
XX
PN US6291662-B1.

XX DNA sequence coding for ompA-tc-met-SLPI.
DE Serine protease inhibitor protein; recombinant; leukocyte elastase;
XX trypsin; secretory leukocyte protease inhibitor; SLPI; ds.

OS Homo sapiens.
XX US6291662-B1.
XX 18-SEP-2001.
XX 22-SEP-1998; 98US-0158085.
XX 30-MAR-1987; 87US-0031846.
XX 06-AUG-1990; 90US-0563832.
XX 22-JUL-1994; 94US-0279056.
XX 05-DEC-1984; 84US-0678222.
XX 02-DEC-1985; 85US-0803471.
XX 29-JUL-1986; 86US-0890526.

PA (AMGE-) AMGEN INC.
XX Bandyopadhyay PK, Eisenberg SP, Stetler GL, Thompson RC;
XX WPI; 2001-637974/73.

XX New DNA sequences, useful in recombinant DNA techniques for directing
PT the production of a serine protease inhibitor protein, e.g. leukocyte
PT elastase or trypsin -

XX Example 3; Column 37-38; 37pp; English.

XX The invention relates to a DNA sequence encoding an analog of a mammalian
CC serine protease inhibitor protein. The analog comprises at least eight
CC cysteine residues and possesses serine protease inhibitor activity, and
CC at least one operational element consisting of a promoter, an operator,
CC a leader sequence, a Shine-Dalgarno sequence, a ribosome binding site
CC or a terminator codon. The DNA is useful in recombinant DNA techniques
CC for directing the production of a serine protease inhibitor protein,
CC e.g. leukocyte elastase or trypsin. The present sequence represents the
CC DNA sequence coding for ompA-tc-met-secretory leukocyte protease
CC inhibitor (SLPI).

XX Sequence 436 BP; 128 A; 102 C; 116 G; 90 T; 0 other;

Query Match 15.6%; Score 32.2; DB 22; Length 436;
Best Local Similarity 51.8%; Pred. No. 0.7;
Matches 73; Conservative 0; Mismatches 68; Indels 0; Gaps 0;

QY 63 AAGCCAGGTTCTTGCTCTATTATCTTGATTCGTTGCGCTATGTTAAACCCACCTAACCGT 122
DB 278 AAACCGGGTAAATGCCCGGTAACTATGCGCAGTGTCTGATGCTGAACCCGCCGAACCTC 337
QY 123 TGTGTTGAAGGACACTGATTGTCCAGGTATCAAAAAGTGTGTAAGGTTCTGCGGTATG 182
DB 338 TCGGAAATGGACGGCCAGTGTAAACGAGATCTGAAATGCTGTATGGGTATGTGCGGCAA 397
QY 183 GCTTGTGTTTCGTTCCACAATAA 203
DB 398 TCTTGTGTTTCCCGGTAAAAA 418

RESULT 39
AAX16271
ID AAX16271 standard; DNA; 460 BP.
XX AAX16271;
AC AAX16271;

XX 19-APR-1999 (first entry)
XX bNA sequence of ompA-SLPI.
DE
XX

KW Serine protease inhibitor; human; emphysema; arthritis; periodontitis;
KW muscular dystrophy; tumour invasion; glomerulonephritis; sepsis;
KW acute leukemia; ds.

XX Homo sapiens.
OS Synthetic.
XX US5871956-A.
XX 16-FEB-1999.

XX 22-JUL-1994; 94US-0279056.
XX 30-MAR-1987; 87US-0031846.
XX 05-DEC-1984; 84US-0678222.
XX 29-JUL-1986; 86US-0890526.
XX 03-SEP-1986; 86US-0903471.
XX 06-AUG-1990; 90US-0563832.
XX 22-JUL-1994; 94US-0279056.

PA (AMGE-) AMGEN INC.
XX Bandyopadhyay PK, Eisenberg SP, Stetler GL, Thompson RC;
XX WPI; 1999-166640/14.

XX New DNA sequence encoding mammalian serine protease inhibitor -
PT useful for recombinantly producing inhibitors with different
PT specificities and treating diseases such as emphysema, arthritis,
PT muscular dystrophy, and tumour invasion

XX Example 3; Column 37-38; 37pp; English.

XX The present invention describes a DNA sequence (A) encoding an analogue
CC of a mammalian serine protease inhibitor (B). The DNA sequences and
CC recombinant methods allow manufacture of a class of inhibitors of e.g.
CC cathepsin G, elastase, and trypsin, with different specificities. The
CC recombinant serine protease product can be directed to act
CC intracellularly or extracellularly and is useful in treating conditions
CC caused by a disturbance in the native protease/protease inhibitor
CC balance such as emphysema, arthritis, glomerulonephritis, periodontitis,
CC muscular dystrophy, tumour invasion, sepsis, and acute leukemia. (A)
CC allows (B) to be recombinantly produced in sufficient quantities and
CC purities so as to provide economical pharmaceutical compositions. (B) is
CC resistant to heat, acid, and a variety of proteolytic enzymes, is
CC thermodynamically stable in extracellular conditions, and exhibits a
CC high degree of self assembly forming an active tertiary structure in the
CC absence of biochemical stimuli. The present sequence represents the DNA
CC sequence of ompA-SLPI.

XX Sequence 460 BP; 123 A; 114 C; 125 G; 98 T; 0 other;

Query Match 15.6%; Score 32.2; DB 20; Length 460;
Best Local Similarity 51.8%; Pred. No. 0.72;
Matches 73; Conservative 0; Mismatches 68; Indels 0; Gaps 0;

QY 63 AAGCCAGGTTCTTGCTCTATTATCTTGATTCGTTGCGCTATGTTAAACCCACCTAACCGT 122
DB 302 AAACCGGGTAAATGCCCGGTAACTATGCGCAGTGTCTGATGCTGAACCCGCCGAACCTC 361
QY 123 TGTGTTGAAGGACACTGATTGTCCAGGTATCAAAAAGTGTGTAAGGTTCTGCGGTATG 182
DB 362 TCGGAAATGGACGGCCAGTGTAAACGAGATCTGAAATGCTGTATGGGTATGTGCGGCAA 421
QY 183 GCTTGTGTTTCGTTCCACAATAA 203
DB 422 TCTTGTGTTTCCCGGTAAAA 442

RESULT 40
AAC97622
ID AAC97622 standard; DNA; 460 BP.
XX

AC AAC97622;
XX
DT 27-FEB-2001 (first entry)
XX
DE DNA encoding OmpA secretory leukocyte protease inhibitor.
XX
KW Serine protease inhibitor; cytostatic; anti-inflammatory; arthritis;
KW protease mediated tissue destruction; emphysema; glomerulonephritis;
KW periodontitis; muscular dystrophy; tumour invasion; chymotrypsin;
KW elastase; ds.
XX
OS Synthetic.
XX
XX US6132990-A.
PN
XX 17-OCT-2000.
PD
XX
XX 07-JUN-1991; 91US-0712354.
PF
XX
XX 03-JAN-1989; 89US-0293042.
PR
XX 06-DEC-1984; 84US-0678822.
PR
XX 02-DEC-1985; 85US-0803471.
PR
XX 29-JUL-1986; 86US-0890526.
PR
XX 30-MAR-1987; 87US-0031846.
PR
XX 04-AUG-1987; 87US-0082962.
XX
PA (AMGE-) AMGEN BOULDER INC.
XX
XX Bandyopadhyay PK, Eisenberg SP, Stetler GL, Thompson RC;
PI
XX WPI; 2000-678667/66.
DR
XX
XX New serine protease inhibitors and DNA sequences for treating a
PT protease-mediated condition or tissue destruction e.g. emphysema or
PT tumor invasion and for recombinant production of inhibitors -
XX
PS Example 3; Column 37-38; 47pp; English.
XX
XX This invention relates to new purified and isolated mammalian serine
CC protease inhibitor proteins which comprise at least 8 cysteine residues
CC and no more than 107 amino acids. The protease inhibitors are capable of
CC inhibiting chymotrypsin and elastase. Sequences AAC97526 - AAC97534,
CC AAC97574 - AAC97581 and AAC97614 - AAC97641 represent DNA encoding the
CC protease inhibitors of the invention, and include oligonucleotide
CC sequences used in the isolation and characterisation of the proteins.
CC Primers AAC97535 - AAC97573 and AAC95582 - AAC97613 are used in the
CC construction of DNA encoding the protease inhibitors. Peptide sequences
CC AAB53098 - AAB53122 represent the protease inhibitor of the invention and
CC various other peptides used in the isolation of the protease inhibitors.
CC The serine protease inhibitors have cytostatic and anti-inflammatory activity.
CC The serine protease inhibitor protein is useful for treating a
CC protease-mediated condition, which includes protease mediated tissue
CC destruction, e.g. emphysema, arthritis, glomerulonephritis,
CC periodontitis, muscular dystrophy or tumour invasion. It is also useful
CC for modulating protease activity. The DNAs are useful for producing the
CC proteins, especially by recombinant methods.
XX
SQ Sequence 460 BP; 123 A; 114 C; 125 G; 98 T; 0 other;

Query Match 15.6%; Score 32.2; DB 21; Length 460;
Best Local Similarity 51.8%; Pred. No. 0.72;
Matches 73; Conservative 0; Mismatches 68; Indels 0; Gaps 0;

QY 63 AAGCCAGGTTCTTGTCTATTATCTTGTATTCGTTGCGCTATGTTAAACCCACCTAACCGT 122
Db 302 AAACCGGGTAAATGCCCGGTAACCTATGGCCAGTGTCTGATGCTGAACCCGCCGAACCTC 361
QY 123 TGTTTGAAGGACACTGATTGTCCAGGTATCAAAAAGTCTGTGAAGTTCCTGCGGTATG 182
Db 362 TCGGAAATGGACGCCAGTGTAAACGAGATCTGAAATGCTGTATGGTATGTGCGGCAA 421
QY 183 GCTTGTTCGTTCCACAATAA 203
Db 422 TCTTGTGTTTCCCGGTAAAA 442

Db 422 TCTTGTGTTTCCCGGTAAAA 442

RESULT 41
AAI67582
ID AAI67582 standard; DNA; 460 BP.
XX
AC AAI67582;
XX
DT 11-FEB-2002 (first entry)
XX
DE DNA sequence coding for ompA SLPI.
XX
KW Serine protease inhibitor protein; recombinant; leukocyte elastase;
KW trypsin; secretory leukocyte protease inhibitor; SLPI; ds.
XX
OS Homo sapiens.
XX
PN US6291662-B1.
XX
PD 18-SEP-2001.
XX
PF 22-SEP-1998; 98US-0158085.
XX
PR 30-MAR-1987; 87US-0031846.
PR 06-AUG-1990; 90US-0563832.
PR 22-JUL-1994; 94US-0279056.
PR 05-DEC-1984; 84US-0678222.
PR 02-DEC-1985; 85US-0803471.
PR 29-JUL-1986; 86US-0890526.
XX
PA (AMGE-) AMGEN INC.
XX
PI Bandyopadhyay PK, Eisenberg SP, Stetler GL, Thompson RC;
XX
XX WPI; 2001-637974/73.
DR
XX
XX New DNA sequences, useful in recombinant DNA techniques for directing
PT the production of a serine protease inhibitor protein, e.g. leukocyte
PT elastase or trypsin -
XX
PS Example 3; Column 36; 37pp; English.
XX
XX The invention relates to a DNA sequence encoding an analog of a mammalian
CC serine protease inhibitor protein. The analog comprises at least eight
CC cysteine residues and possesses serine protease inhibitor activity, and
CC at least one operational element consisting of a promoter, an operator,
CC a leader sequence, a Shine-Dalgarno sequence, a ribosome binding site
CC or a terminator codon. The DNA is useful in recombinant DNA techniques
CC for directing the production of a serine protease inhibitor protein,
CC e.g. leukocyte elastase or trypsin. The present sequence represents the
CC DNA sequence coding for ompA-secretory leukocyte protease inhibitor
CC (SLPI).
XX
SQ Sequence 460 BP; 123 A; 114 C; 125 G; 98 T; 0 other;

Query Match 15.6%; Score 32.2; DB 22; Length 460;
Best Local Similarity 51.8%; Pred. No. 0.72;
Matches 73; Conservative 0; Mismatches 68; Indels 0; Gaps 0;

QY 63 AAGCCAGGTTCTTGTCTATTATCTTGTATTCGTTGCGCTATGTTAAACCCACCTAACCGT 122
Db 302 AAACCGGGTAAATGCCCGGTAACCTATGGCCAGTGTCTGATGCTGAACCCGCCGAACCTC 361
QY 123 TGTTTGAAGGACACTGATTGTCCAGGTATCAAAAAGTCTGTGAAGTTCCTGCGGTATG 182
Db 362 TCGGAAATGGACGCCAGTGTAAACGAGATCTGAAATGCTGTATGGTATGTGCGGCAA 421
QY 183 GCTTGTTCGTTCCACAATAA 203
Db 422 TCTTGTGTTTCCCGGTAAAA 442

```

Db      252 TACCACATACCATGATTCACTCTGTCATCATTAATTCACATAAACATATCCAGGGAATGT 193
QY      122 TTGTTGAAGGACACTGATTGTCCAGGTATCAAAAAGTGCTGTGAAGGTTCTCTGCGGTAT 181
Db      192 TTTTTTAACGGTTGTTTTAGCTTTACCATCTTTAACTTGAGTTTCTTCTTCTCCGGTAT 133
QY      182 GGCTTGTTT 190
Db      132 GACTACTCT 124

RESULT 43
AAS54717/c
ID AAS54717 standard; DNA; 549 BP.
XX
AC AAS54717;
XX
DT 13-FEB-2002 (first entry)
XX
DE Staphylococcus aureus DNA for cellular proliferation protein #1029.
XX
XX Antisense; ds; prokaryotic cellular proliferation gene;
KW antibiotic; antibacterial; drug design.
KW
XX
OS Staphylococcus aureus.
XX
XX WO200170955-A2.
PN
XX
XX 27-SEP-2001.
PD
XX
XX 21-MAR-2001; 2001WO-US09180.
PF
XX
XX 21-MAR-2000; 2000US-191078P.
PR
XX 23-MAY-2000; 2000US-206848P.
PR
XX 26-MAY-2000; 2000US-207727P.
PR
XX 23-OCT-2000; 2000US-242578P.
PR
XX 27-NOV-2000; 2000US-253625P.
PR
XX 22-DEC-2000; 2000US-257931P.
PR
XX 16-FEB-2001; 2001US-269308P.
XX
XX (ELIT-) ELITRA PHARM INC.
PA
XX
XX Haselbeck R, Ohlsen KL, Zyskind JW, Wall D, Trawick JD, Carr GJ;
PI Yamamoto RT, Xu HH;
PI
XX
XX WPI; 2001-611495/70.
DR
XX P-PSDB; AAU36858.
DR
XX
XX New polynucleotides for the identification and development of
PT antibiotics, comprise sequences of antisense nucleic acids -
PT
XX
XX Claim 27; Seq ID No 8354; 511pp; English.
PS
XX
XX The invention relates to antisense inhibitors of genes essential to
CC prokaryotic cellular proliferation, their use in identifying the
CC genes, their use in the discovery of novel antibiotics, the essential
CC genes themselves and the encoded proteins. The prokaryotes used are
CC Escherichia coli, Staphylococcus aureus, Salmonella typhi, Klebsiella
CC pneumoniae, Pseudomonas aeruginosa and Enterococcus faecalis. The
CC invention is also useful for the identification of potential new targets
CC for antibiotic development. The antisense nucleic acids can also be used
CC to identify proteins used in proliferation, to express these proteins,
CC and to obtain antibodies capable of binding to the expressed proteins.
CC The proteins can be used to screen compounds in rational drug discovery
CC programmes. The antisense nucleic acid sequence is also useful to screen
CC for homologous nucleic acids which are required for cell proliferation in
CC a wide variety of organisms. The present sequence encodes an
CC essential prokaryotic cellular proliferation protein.
CC
CC Note: The sequence data for this patent did not form part
CC of the printed specification, but was obtained in electronic
CC format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences.

```

```

XX
SQ Sequence 549 BP; 207 A; 73 C; 122 G; 147 T; 0 other;
Query Match 15.6%; Score 32.2; DB 23; Length 549;
Best Local Similarity 48.1%; Pred. No. 0.77;
Matches 91; Conservative 0; Mismatches 98; Indels 0; Gaps 0;

QY 2 ATTCGAGCTCGGTACCATACATGCTCAAGAACCCAGTTAAAGGTCCTGTCTAC 61
Db 312 ATTTGGCTTAGACCCCTGCACCTGCAGAACCTACAAAACCCAGTAACGCCCTGGTGATTTCT 253

QY 62 TAAGCCAGGTTCTTGTCTCTATTATCTTGATTCGTTGCGCTATGTTAAACCCACCTAACCG 121
Db 252 TACCACATACCATGATTCATCTGTCATGATTAATCCACTAAACATATCCAGGGAATGT 193

QY 122 TTGTTTGAAGGACACTGATGTCAGGATCAAAAAGTGCTGTGAAGGTTCTTCGCGGTAT 181
Db 192 TTTTAAACGGTTGTTTAGCTTTACCATCTTTAACTTGAGTTTCTTCTTCTTCCGGTAT 133

QY 182 GGCTTGTTT 190
Db 132 GACTACTCT 124

RESULT 44
AAK65489
ID AAK65489 standard; DNA; 16605 BP.
XX
AC AAK65489;
XX
DT 06-NOV-2001 (first entry)
XX
DE Human immune/haematopoietic antigen genomic sequence SEQ ID NO:20301.
XX
KW Human; immune; haematopoietic; immune/haematopoietic antigen; cancer;
KW cytostatic; gene therapy; vaccine; metastasis; ds.
XX
OS Homo sapiens.
XX
PN WO200157182-A2.
XX
PD 09-AUG-2001.
XX
PF 17-JAN-2001; 2001WO-US01354.
XX
PR 31-JAN-2000; 2000US-0179065.
PR 04-FEB-2000; 2000US-0180628.
PR 24-FEB-2000; 2000US-0184664.
PR 02-MAR-2000; 2000US-0186350.
PR 16-MAR-2000; 2000US-0189874.
PR 17-MAR-2000; 2000US-0190076.
PR 18-APR-2000; 2000US-0198123.
PR 19-MAY-2000; 2000US-0205515.
PR 07-JUN-2000; 2000US-0209467.
PR 28-JUN-2000; 2000US-0214886.
PR 30-JUN-2000; 2000US-0215135.
PR 07-JUL-2000; 2000US-0216647.
PR 07-JUL-2000; 2000US-0216880.
PR 11-JUL-2000; 2000US-0217487.
PR 14-JUL-2000; 2000US-0217496.
PR 26-JUL-2000; 2000US-0218290.
PR 26-JUL-2000; 2000US-0220963.
PR 14-AUG-2000; 2000US-0220964.
PR 14-AUG-2000; 2000US-0224518.
PR 14-AUG-2000; 2000US-0224519.
PR 14-AUG-2000; 2000US-0225213.
PR 14-AUG-2000; 2000US-0225214.
PR 14-AUG-2000; 2000US-0225266.
PR 14-AUG-2000; 2000US-0225267.
PR 14-AUG-2000; 2000US-0225268.
PR 14-AUG-2000; 2000US-0225270.
PR 14-AUG-2000; 2000US-0225447.
PR 14-AUG-2000; 2000US-0225757.
PR 14-AUG-2000; 2000US-0225758.
PR 14-AUG-2000; 2000US-0225759.
PR 18-AUG-2000; 2000US-0226279.
PR 22-AUG-2000; 2000US-0226681.
PR 22-AUG-2000; 2000US-0226868.
PR 22-AUG-2000; 2000US-0227182.
PR 23-AUG-2000; 2000US-0227009.
PR 30-AUG-2000; 2000US-0228924.
PR 01-SEP-2000; 2000US-0229287.
PR 01-SEP-2000; 2000US-0229343.
PR 01-SEP-2000; 2000US-0229344.
PR 01-SEP-2000; 2000US-0229345.
PR 05-SEP-2000; 2000US-0229509.
PR 05-SEP-2000; 2000US-0229513.
PR 06-SEP-2000; 2000US-0230437.
PR 06-SEP-2000; 2000US-0230438.
PR 08-SEP-2000; 2000US-0231242.
PR 08-SEP-2000; 2000US-0231243.
PR 08-SEP-2000; 2000US-0231244.
PR 08-SEP-2000; 2000US-0231413.
PR 08-SEP-2000; 2000US-0231414.
PR 08-SEP-2000; 2000US-0232080.
PR 08-SEP-2000; 2000US-0232081.
PR 12-SEP-2000; 2000US-0231968.
PR 14-SEP-2000; 2000US-0232397.
PR 14-SEP-2000; 2000US-0232398.
PR 14-SEP-2000; 2000US-0232399.
PR 14-SEP-2000; 2000US-0232400.
PR 14-SEP-2000; 2000US-0232401.
PR 14-SEP-2000; 2000US-0233063.
PR 14-SEP-2000; 2000US-0233064.
PR 14-SEP-2000; 2000US-0233065.
PR 21-SEP-2000; 2000US-0234223.
PR 21-SEP-2000; 2000US-0234274.
PR 25-SEP-2000; 2000US-0234997.
PR 25-SEP-2000; 2000US-0234998.
PR 26-SEP-2000; 2000US-0235484.
PR 27-SEP-2000; 2000US-0235834.
PR 27-SEP-2000; 2000US-0235836.
PR 29-SEP-2000; 2000US-0236327.
PR 29-SEP-2000; 2000US-0236367.
PR 29-SEP-2000; 2000US-0236368.
PR 29-SEP-2000; 2000US-0236369.
PR 29-SEP-2000; 2000US-0236370.
PR 02-OCT-2000; 2000US-0236802.
PR 02-OCT-2000; 2000US-0237037.
PR 02-OCT-2000; 2000US-0237038.
PR 02-OCT-2000; 2000US-0237039.
PR 02-OCT-2000; 2000US-0237040.
PR 13-OCT-2000; 2000US-0239935.
PR 13-OCT-2000; 2000US-0239937.
PR 20-OCT-2000; 2000US-0240960.
PR 20-OCT-2000; 2000US-0241221.
PR 20-OCT-2000; 2000US-0241785.
PR 20-OCT-2000; 2000US-0241786.
PR 20-OCT-2000; 2000US-0241787.
PR 20-OCT-2000; 2000US-0241808.
PR 20-OCT-2000; 2000US-0241809.
PR 01-NOV-2000; 2000US-0241826.
PR 08-NOV-2000; 2000US-0244617.
PR 08-NOV-2000; 2000US-0246474.
PR 08-NOV-2000; 2000US-0246475.
PR 08-NOV-2000; 2000US-0246476.
PR 08-NOV-2000; 2000US-0246477.
PR 08-NOV-2000; 2000US-0246478.
PR 08-NOV-2000; 2000US-0246523.
PR 08-NOV-2000; 2000US-0246524.
PR 08-NOV-2000; 2000US-0246525.
PR 08-NOV-2000; 2000US-0246526.
PR 08-NOV-2000; 2000US-0246527.
PR 08-NOV-2000; 2000US-0246528.
PR 08-NOV-2000; 2000US-0246532.
PR 08-NOV-2000; 2000US-0246609.
```

Sun Feb 16 09:12:52 2003

PR 08-NOV-2000; 2000US-0246610.
PR 08-NOV-2000; 2000US-0246611.
PR 08-NOV-2000; 2000US-0246613.
PR 17-NOV-2000; 2000US-0249207.
PR 17-NOV-2000; 2000US-0249208.
PR 17-NOV-2000; 2000US-0249209.
PR 17-NOV-2000; 2000US-0249210.
PR 17-NOV-2000; 2000US-0249211.
PR 17-NOV-2000; 2000US-0249212.
PR 17-NOV-2000; 2000US-0249213.
PR 17-NOV-2000; 2000US-0249214.
PR 17-NOV-2000; 2000US-0249215.
PR 17-NOV-2000; 2000US-0249216.
PR 17-NOV-2000; 2000US-0249217.
PR 17-NOV-2000; 2000US-0249218.
PR 17-NOV-2000; 2000US-0249244.
PR 17-NOV-2000; 2000US-0249245.
PR 17-NOV-2000; 2000US-0249264.
PR 17-NOV-2000; 2000US-0249265.
PR 17-NOV-2000; 2000US-0249297.
PR 17-NOV-2000; 2000US-0249299.
PR 17-NOV-2000; 2000US-0249300.
PR 01-DEC-2000; 2000US-0250160.
PR 01-DEC-2000; 2000US-0250391.
PR 05-DEC-2000; 2000US-0251030.
PR 05-DEC-2000; 2000US-0251988.
PR 05-DEC-2000; 2000US-0256719.
PR 06-DEC-2000; 2000US-0251479.
PR 08-DEC-2000; 2000US-0251856.
PR 08-DEC-2000; 2000US-0251868.
PR 08-DEC-2000; 2000US-0251869.
PR 08-DEC-2000; 2000US-0251989.
PR 08-DEC-2000; 2000US-0251990.
PR 11-DEC-2000; 2000US-0254097.
PR 05-JAN-2001; 2001US-0259678.

(HUMA-) HUMAN GENOME SCI INC.

Rosen CA, Barash SC, Ruben SM;

WPI; 2001-483426/52.

Nucleic acids encoding human immune/hematopoietic antigen polypeptides,
useful for preventing, diagnosing and/or treating cancers and
metastasis -

Disclosure; SEQ ID NO 20301; 3071pp + Sequence Listing; English.

AAK54951 to AAK64702 encode the human immune/haematopoietic antigen (I)
amino acid sequences given in AAM82170 to AAM91921. (I) have cytostatic
activity, and can be used in gene therapy and vaccine production. (I)
proteins and polynucleotides may be used in the prevention, diagnosis and
treatment of diseases associated with inappropriate (I) expression. For
example, they may be used to treat disorders associated with decreased
expression by rectifying mutations or deletions in a patient's genome
that affect the activity of (I) by expressing inactive proteins or to
supplement the patients own production of (I). Additionally, (I)
polynucleotides may be used to produce the secreted (I), by inserting
the nucleic acids into a host cell and culturing the cell to express the
protein. (I) proteins and polynucleotides may be used to prevent,
diagnose and treat immune/haematopoietic-related diseases, especially
cancers and cancer metastases of haematopoietic-derived cells. AAK64703
to AAK87694 represent human immune/haematopoietic antigen genomic
sequences from the present invention. AAK54942 to AAK54950 and AAM82169
represent sequences used in the exemplification of the present invention.

Sequence 16605 BP; 4009 A; 3696 C; 4212 G; 4688 T; 0 other;

Query Match 15.3%; Score 31.6; DB 22; Length 16605;
Best Local Similarity 52.2%; Pred. NO. 4.4;
Matches 70; Conservative 0; Mismatches 64; Indels 0; Gaps 0;

QY 55 TGCTACTAAGCCAGGTTCTTCTGCTATTATCTTGATTCGTTGCGTATGTTAAACCCAC 114

Db 4004 TTTCTATTATGTCCTTCTTTAGGCCATTATTTCTGACCCCTGGGGCAGGATCACTCACT 4063
QY 115 CTAACCGTTGTTTGAAGGACACTGATTGTCCAGGTATCAAAAAGTCTGTGAAGGTTTCCT 174
Db 4064 TGGGGCCTTATAGCTGGGACACTGACGCTCAAAATACCAGGAGCTGCTGGAATGGGTATT 4123
QY 175 GCGGTATGGCTTGT 188
Db 4124 GTAATATGTATGGT 4137
RESULT 45
AAK87049
ID AAK87049 standard; DNA; 16605 BP.
XX
AC AAK87049;
XX 07-NOV-2001 (first entry)
DT Human immune/haematopoietic antigen genomic sequence SEQ ID NO:41861.
XX Human; immune; haematopoietic; immune/haematopoietic antigen; cancer;
KW cytostatic; gene therapy; vaccine; metastasis; ds.
XX Homo sapiens.
OS
XX WO200157182-A2.
PN
XX 09-AUG-2001.
PD
PF 17-JAN-2001; 2001WO-US01354.
XX
PR 31-JAN-2000; 2000US-0179065.
PR 04-FEB-2000; 2000US-0180628.
PR 24-FEB-2000; 2000US-0184664.
PR 02-MAR-2000; 2000US-0186350.
PR 16-MAR-2000; 2000US-0189874.
PR 17-MAR-2000; 2000US-0190076.
PR 18-APR-2000; 2000US-0198123.
PR 19-MAY-2000; 2000US-0205515.
PR 07-JUN-2000; 2000US-0209467.
PR 28-JUN-2000; 2000US-0214886.
PR 30-JUN-2000; 2000US-0215135.
PR 07-JUL-2000; 2000US-0216647.
PR 07-JUL-2000; 2000US-0216880.
PR 11-JUL-2000; 2000US-0217487.
PR 11-JUL-2000; 2000US-0217496.
PR 14-JUL-2000; 2000US-0218290.
PR 26-JUL-2000; 2000US-0220963.
PR 26-JUL-2000; 2000US-0220964.
PR 14-AUG-2000; 2000US-0224518.
PR 14-AUG-2000; 2000US-0224519.
PR 14-AUG-2000; 2000US-0225213.
PR 14-AUG-2000; 2000US-0225214.
PR 14-AUG-2000; 2000US-0225267.
PR 14-AUG-2000; 2000US-0225267.
PR 14-AUG-2000; 2000US-0225268.
PR 14-AUG-2000; 2000US-0225270.
PR 14-AUG-2000; 2000US-0225477.
PR 14-AUG-2000; 2000US-0225757.
PR 14-AUG-2000; 2000US-0225758.
PR 14-AUG-2000; 2000US-0225759.
PR 18-AUG-2000; 2000US-0226279.
PR 22-AUG-2000; 2000US-0226681.
PR 22-AUG-2000; 2000US-0226688.
PR 22-AUG-2000; 2000US-0227182.
PR 23-AUG-2000; 2000US-0227009.
PR 30-AUG-2000; 2000US-0228924.
PR 01-SEP-2000; 2000US-0229287.
PR 01-SEP-2000; 2000US-0229343.
PR 01-SEP-2000; 2000US-0229344.
PR 01-SEP-2000; 2000US-0229345.

AB18664	standard; DNA; 3496 BP.
AB18664	
23-JAN-2002	(first entry)
Human nervous system related polynucleotide SEQ ID NO 10995.	
Human; neurotropic; neuroprotective; cytostatic; dermatological; virucide; immunosuppressive; antiinflammatory; anti-HIV; antibacterial; vulnerary; antiparkinsonian; antistickling; antianaemic; antiarthritic; cancer; antirheumatic; hepatotropic; cerebroprotective; antiinflammatory; anti allergic; antidiabetic; antiulcer; anticonvulsant; antifungal; antiparasitic; cardiac; immune disorder; cardiovascular disorder; neurological disease; infection; nephrotropic; gene therapy; vaccine; ds	
Homo sapiens.	
WO200159063-A2.	
16-AUG-2001.	
17-JAN-2001; 2001WO-US01334.	
31-JAN-2000; 2000US-0179065.	
04-FEB-2000; 2000US-0180628.	
24-FEB-2000; 2000US-0184664.	
02-MAR-2000; 2000US-0186350.	
16-MAR-2000; 2000US-0189874.	
17-MAR-2000; 2000US-0190076.	
18-APR-2000; 2000US-0198123.	
19-MAY-2000; 2000US-0205515.	
07-JUN-2000; 2000US-0209467.	
28-JUN-2000; 2000US-0214886.	
30-JUN-2000; 2000US-0215135.	
07-JUL-2000; 2000US-0216647.	
07-JUL-2000; 2000US-0216880.	
11-JUL-2000; 2000US-0217487.	
11-JUL-2000; 2000US-0217496.	
14-JUL-2000; 2000US-0218290.	
26-JUL-2000; 2000US-0220963.	
26-JUL-2000; 2000US-0220964.	
14-AUG-2000; 2000US-0224518.	
14-AUG-2000; 2000US-0224519.	
14-AUG-2000; 2000US-0225213.	
14-AUG-2000; 2000US-0225214.	
14-AUG-2000; 2000US-0225266.	
14-AUG-2000; 2000US-0225267.	
14-AUG-2000; 2000US-0225268.	
14-AUG-2000; 2000US-0225270.	
14-AUG-2000; 2000US-0225447.	
14-AUG-2000; 2000US-0225757.	
14-AUG-2000; 2000US-0225758.	
14-AUG-2000; 2000US-0225759.	
18-AUG-2000; 2000US-0226279.	
22-AUG-2000; 2000US-0226681.	
22-AUG-2000; 2000US-0226868.	
22-AUG-2000; 2000US-0227182.	
23-AUG-2000; 2000US-0227009.	
30-AUG-2000; 2000US-0228924.	
01-SEP-2000; 2000US-0229287.	
01-SEP-2000; 2000US-0229343.	
01-SEP-2000; 2000US-0229344.	
01-SEP-2000; 2000US-0229345.	
05-SEP-2000; 2000US-0229509.	
05-SEP-2000; 2000US-0229513.	
06-SEP-2000; 2000US-0230437.	
06-SEP-2000; 2000US-0230438.	
08-SEP-2000; 2000US-0231242.	
08-SEP-2000; 2000US-0231243.	
08-SEP-2000; 2000US-0231244.	
08-SEP-2000; 2000US-0231413.	
08-SEP-2000; 2000US-0231414.	

PR	08-SEP-2000;	2000US-0232080;
PR	08-SEP-2000;	2000US-0232081;
PR	12-SEP-2000;	2000US-0231968;
PR	14-SEP-2000;	2000US-0232397;
PR	14-SEP-2000;	2000US-0232398;
PR	14-SEP-2000;	2000US-0232399;
PR	14-SEP-2000;	2000US-0232400;
PR	14-SEP-2000;	2000US-0232401;
PR	14-SEP-2000;	2000US-0233063;
PR	14-SEP-2000;	2000US-0233064;
PR	14-SEP-2000;	2000US-0233065;
PR	21-SEP-2000;	2000US-0234223;
PR	21-SEP-2000;	2000US-0234274;
PR	25-SEP-2000;	2000US-0234997;
PR	25-SEP-2000;	2000US-0234998;
PR	26-SEP-2000;	2000US-0235484;
PR	27-SEP-2000;	2000US-0235834;
PR	27-SEP-2000;	2000US-0235836;
PR	29-SEP-2000;	2000US-0236327;
PR	29-SEP-2000;	2000US-0236367;
PR	29-SEP-2000;	2000US-0236368;
PR	29-SEP-2000;	2000US-0236369;
PR	29-SEP-2000;	2000US-0236370;
PR	02-OCT-2000;	2000US-0236802;
PR	02-OCT-2000;	2000US-0237037;
PR	02-OCT-2000;	2000US-0237038;
PR	02-OCT-2000;	2000US-0237039;
PR	02-OCT-2000;	2000US-0237040;
PR	13-OCT-2000;	2000US-0239935;
PR	13-OCT-2000;	2000US-0239937;
PR	20-OCT-2000;	2000US-0240960;
PR	20-OCT-2000;	2000US-0241785;
PR	20-OCT-2000;	2000US-0241786;
PR	20-OCT-2000;	2000US-0241787;
PR	20-OCT-2000;	2000US-0241808;
PR	20-OCT-2000;	2000US-0241809;
PR	20-OCT-2000;	2000US-0241826;
PR	20-OCT-2000;	2000US-0242221;
PR	01-NOV-2000;	2000US-0244617;
PR	08-NOV-2000;	2000US-0246474;
PR	08-NOV-2000;	2000US-0246475;
PR	08-NOV-2000;	2000US-0246476;
PR	08-NOV-2000;	2000US-0246477;
PR	08-NOV-2000;	2000US-0246478;
PR	08-NOV-2000;	2000US-0246523;
PR	08-NOV-2000;	2000US-0246524;
PR	08-NOV-2000;	2000US-0246525;
PR	08-NOV-2000;	2000US-0246526;
PR	08-NOV-2000;	2000US-0246527;
PR	08-NOV-2000;	2000US-0246528;
PR	08-NOV-2000;	2000US-0246532;
PR	08-NOV-2000;	2000US-0246609;
PR	08-NOV-2000;	2000US-0246610;
PR	08-NOV-2000;	2000US-0246611;
PR	08-NOV-2000;	2000US-0246613;
PR	17-NOV-2000;	2000US-0249207;
PR	17-NOV-2000;	2000US-0249208;
PR	17-NOV-2000;	2000US-0249209;
PR	17-NOV-2000;	2000US-0249211;
PR	17-NOV-2000;	2000US-0249212;
PR	17-NOV-2000;	2000US-0249213;
PR	17-NOV-2000;	2000US-0249214;
PR	17-NOV-2000;	2000US-0249215;
PR	17-NOV-2000;	2000US-0249216;
PR	17-NOV-2000;	2000US-0249217;
PR	17-NOV-2000;	2000US-0249218;
PR	17-NOV-2000;	2000US-0249244;
PR	17-NOV-2000;	2000US-0249245;
PR	17-NOV-2000;	2000US-0249264;
PR	17-NOV-2000;	2000US-0249265;
PR	17-NOV-2000;	2000US-0249297;
PR	17-NOV-2000;	2000US-0249299;

PR 17-NOV-2000; 2000US-0249300.
 PR 01-DEC-2000; 2000US-0250391.
 PR 01-DEC-2000; 2000US-0251160.
 PR 05-DEC-2000; 2000US-0251030.
 PR 05-DEC-2000; 2000US-0251988.
 PR 05-DEC-2000; 2000US-0256719.
 PR 06-DEC-2000; 2000US-0251479.
 PR 08-DEC-2000; 2000US-0251856.
 PR 08-DEC-2000; 2000US-0251868.
 PR 08-DEC-2000; 2000US-0251869.
 PR 08-DEC-2000; 2000US-0251989.
 PR 08-DEC-2000; 2000US-0251990.
 PR 11-DEC-2000; 2000US-0254097.
 PR 05-JAN-2001; 2001US-0259678.
 XX
 PA (HUMA-) HUMAN GENOME SCI INC.
 XX
 PI Rosen CA, Barash SC, Ruben SM;
 XX
 DR WPI; 2001-541565/60.
 XX
 PT Nucleic acids encoding 3224 human nervous system antigen polypeptides,
 PT useful for preventing, diagnosing and/or treating nervous system
 PT cancers and metastases -
 XX
 PS Disclosure; SEQ ID NO 10995; 1701pp + Sequence Listing; English.
 XX
 CC The invention relates to novel genes (ABA11004-ABA21534) and proteins
 CC (ABB14678-ABB18001) useful for preventing, treating or ameliorating
 CC medical conditions e.g. by protein or gene therapy. The genes are
 CC isolated from a range of human tissues disclosed in the specification.
 CC The nucleic acids, proteins, antibodies and (ant)agonists are useful
 CC in the diagnosis, treatment and prevention of: (a) cancer, e.g. breast
 CC and ovarian cancer and other cancers of the adrenal gland, bone, bone
 CC marrow, breast, gastrointestinal tract, liver, lung, or urogenital;
 CC (b) immune disorders e.g. Addison's disease, allergies, autoimmune
 CC haemolytic anaemia, autoimmune thyroiditis, diabetes mellitus, Crohn's
 CC disease, multiple sclerosis, rheumatoid arthritis and ulcerative
 CC colitis; (c) cardiovascular disorders such as myocardial ischaemias;
 CC (d) wound healing; (e) neurological diseases e.g. cerebral anoxia and
 CC epilepsy; and (f) infectious diseases such as viral, bacterial, fungal
 CC and parasitic infections.
 CC Note: The sequence data for this patent did not form part of the
 CC printed specification, but was obtained in electronic format directly
 CC from WIPO at ftp.wipo.int/pub/published_pct_sequences.
 XX
 SQ Sequence 3496 BP; 1089 A; 587 C; 579 G; 1241 T; 0 other;

[illegible]

RESULT 48
ABQ88179
ID ABQ88179 standard; cDNA; 160771 BP
XX
AC ABQ88179;
XX
DT 18-SEP-2002 (first entry)
XX


```
FT FT 39242..39904
FT FT /*tag= j
FT FT /label= MG033
FT FT /note= "Previously identified as MORF-20100, the
FT FT encoded protein shows 35.90 percentage
FT FT identity to glycerol uptake facilitator
FT FT (glpF) from B. subtilis"
FT FT complement (39873..40514)
FT FT /*tag= k
FT FT /label= MG034
FT FT /note= "Previously identified as MORF-20101, the
FT FT encoded protein shows 48.13 percentage
FT FT identity to thymidylate kinase (tdk)
FT FT from B. subtilis"
FT FT 40543..41787
FT FT /*tag= l
FT FT /label= MG035
FT FT /note= "Previously identified as MORF-20102, the
FT FT encoded protein shows 30.71 percentage
FT FT identity to histidyl-tRNA synthetase (hiss)
FT FT from Mycobacterium leprae"
FT FT complement (44751..46277)
FT FT /*tag= m
FT FT /label= MG038
FT FT /note= "Previously identified as MORF-20105, the
FT FT encoded protein shows 46.83 percentage
FT FT identity to glycerol kinase (glpK)
FT FT from E. coli"
FT FT complement (46268..47422)
FT FT /*tag= n
FT FT /label= MG039
FT FT /note= "Previously identified as MORF-19831 and
FT FT MORF-20106, the encoded protein shows 43.20
FT FT percentage identity to glycerol-3-phosphate
FT FT dehydrogenase (GUT2) from S. cerevisiae"
FT FT 49377..49643
FT FT /*tag= o
FT FT /label= MG041
FT FT /note= "The encoded protein shows 48.86 percentage
FT FT identity to phosphohistidinoprotein-hexose
FT FT phosphotransferase (ptsH) from Mycoplasma
FT FT capricolum"
FT FT 50060..51520
FT FT /*tag= p
FT FT /label= MG042
FT FT /note= "Previously identified as MORF-19832 and
FT FT MORF-20108, the encoded protein shows 41.92
FT FT percentage identity to spermidine/
FT FT putrescine transport ATP-binding protein
FT FT (potA) from E. coli"
FT FT 51525..52382
FT FT /*tag= q
FT FT /label= MG043
FT FT /note= "Previously identified as MORF-20110, the
FT FT encoded protein shows 26.51 percentage
FT FT identity to spermidine/putrescine transport
FT FT system permease protein (potB) from E. coli"
FT FT 52366..53220
FT FT /*tag= r
FT FT /label= MG044
FT FT /note= "Previously identified as MORF-20111, the
FT FT encoded protein shows 29.45 percentage
FT FT identity to spermidine/putrescine transport
FT FT system permease protein C (potC) from E. coli"
FT FT 54658..55605
FT FT /*tag= s
FT FT /label= MG046
FT FT /note= "Previously identified as MORF-20112, the
FT FT encoded protein shows 36.60 percentage
FT FT identity to sialoglycoprotease (gcp)
FT FT from Pasteurella haemolytica"
FT FT complement (56970..58310)
FT FT /*tag= t
```

```
FT FT /label= MG048
FT FT /note= "Previously identified as MORF-19834,
FT FT MORF-20114 and MORF-20115, the encoded protein
FT FT shows 43.02 percentage identity to signal
FT FT recognition particle protein (ffh) from B.
FT FT subtilis"
FT FT 58117..59079
FT FT /*tag= u
FT FT /label= MG049
FT FT /note= "Previously identified as MORF-20114 and
FT FT MORF-20115, the encoded protein shows 44.78
FT FT percentage identity to purine-nucleoside
FT FT phosphorylase (deod) from E. coli"
FT FT 59083..59754
FT FT /*tag= v
FT FT /label= MG050
FT FT /note= "Previously identified as MORF-20117, the
FT FT encoded protein shows 83.03 percentage
FT FT identity to deoxyribose-phosphate aldolase
FT FT (deoc) from Mycoplasma pneumoniae"
FT FT complement (64898..65731)
FT FT /*tag= w
FT FT /label= MG056
FT FT /note= "Previously identified as MORF-20122, the
FT FT encoded protein shows 30.25 percent
FT FT identity to the protein disclosed in
FT FT GB:D26185_99 from B. subtilis"
FT FT complement (65713..66249)
FT FT /*tag= x
FT FT /label= MG057
FT FT /note= "Previously identified as MORF-20123, the
FT FT encoded protein shows 38.90 percentage
FT FT identity to the protein disclosed in
FT FT GB:D26185_104 from B. subtilis"
FT FT 81047..82597
FT FT /*tag= y
FT FT /label= MG067
FT FT /note= "Previously identified as MORF-19845, the
FT FT encoded protein shows 28.84 percentage
FT FT identity to glutamic acid specific protease
FT FT (SPase) from Staphylococcus aureus"
FT FT 91065..91919
FT FT /*tag= z
FT FT /label= MG070
FT FT /note= "Previously identified as MORF-20136, the
FT FT encoded protein shows 34.8 percentage
FT FT identity to ribosomal protein S2 (rps2)
FT FT from Spirulina plantensis"

Query Match 14.8%; Score 30.4; DB 18; Length 580073;
Best Local Similarity 53.3%; Pred. No. 41;
Matches 64; Conservative 0; Mismatches 56; Indels 0; Gaps 0;

QY 37 AACCAAGTTAAAGTCCTGTCTACTAAGCCAGGTTCTTGTCCTATATCTTGATTCGTT 96
Db 276182 AACTAGTTAAGGAAACTGAAATTAGAAACCAATGGTTCTCCTTTTATGCTATTTT 276241
QY 97 GCGCTATGTTAAACCCACCTAACCGTTGTTGAAGGACACTGATGTCAGGTATCAAAA 156
Db 276242 CATCTATCTCCCAACACACCCACCGTTTATTGAAAAAAGCTGGGTTTCAACACAAAAGAA 276301

RESULT 50
AAC07910
ID AAC07910 standard; cDNA; 450 BP.
XX
AC AAC07910;
XX
DT 06-OCT-2000 (first entry)
XX
DE Human secreted protein 5' EST, SEQ ID NO: 11985.
XX
KW Human; 5' EST; expressed sequence tag; secreted protein; cDNA isolation;
```

KW gene therapy; chromosome mapping; ss.
XX
OS Homo sapiens.
XX
PN EP1033401-A2.
XX
PD 06-SEP-2000.
XX
PF 21-FEB-2000; 2000EP-0200610.
XX
PR 26-FEB-1999; 99US-0122487.
XX
XX (GEST) GENSET.
XX
PI Dumas Milne Edwards J, Duclert A, Giordano J;
XX
XX WPI; 2000-500381/45.
DR
XX
XX New nucleic acid that is a 5' expressed sequence tag (5' EST) for
PT obtaining cDNAs and genomic DNAs that correspond to 5'ESTs and for
PT diagnostic, forensic, gene therapy and chromosome mapping procedures -
XX
XX Claim 1; SEQ ID 11985; 71pp + CD-ROM; English.
PS
XX The present sequence is one of a large number of 5' ESTs derived from
CC mRNAs encoding secreted proteins. No ORF has yet been conclusively
CC identified within the present sequence. The 5' ESTs were prepared from
CC total human RNAs or polyA+ RNAs derived from 30 different tissues. EST
CC sequences usually correspond mainly to the 3' untranslated region (UTR)
CC of the mRNA because they are often obtained from oligo-dT primed cDNA
CC libraries. Such ESTs are not well suited for isolating cDNA sequences
CC derived from the 5' ends of mRNAs and even in those cases where longer
CC cDNA sequences have been obtained, the full 5' UTR is rarely included.
CC 5' ESTs are derived from mRNAs with intact 5' ends and can therefore be
CC used to obtain full length cDNAs and genomic DNAs. 5' ESTs are also used
CC in diagnostic, forensic, gene therapy and chromosome mapping procedures.
CC They are used to obtain upstream regulatory sequences and to design
CC expression and secretion vectors.
XX
SQ Sequence 450 BP; 114 A; 89 C; 103 G; 140 T; 4 other;

Query Match 14.7%; Score 30.2; DB 21; Length 450;
Best Local Similarity 50.3%; Pred. No. 3.4;
Matches 74; Conservative 0; Mismatches 73; Indels 0; Gaps 0;

QY 35 AGAACCCAGTTAAAGGTCCTGTCTACTAAGCCAGGTTCTTGTCTTATTATCTTGATTTCG 94
Db 147 AGAAAGGTTAAAGGCTGATGGTACCTTAAGCCTGGTACTTGAATTTGATCAAGATAAG 206
QY 95 TTGCGCTATGTTAAACCCACCTAACCGTTGTTGAAGGACACTGATTGTCAGGTATCAA 154
Db 207 CTGCCTTAAGTTCTCTTCATTACACAAATGATCCTAGATAATTGATAGATCCTGTGGTTC 266
QY 155 AAAGTGCTGTGAAGGTTCTCCTGCGGTAT 181
Db 267 AACTGGATTYYTAGATAGAAGCTGGAT 293

RESULT 51
AAH64937
ID AAH64937 standard; cDNA; 1489 BP.
XX
AC AAH64937;
XX
DT 11-SEP-2001 (first entry)
XX
DE Human secreted protein cDNA, SEQ ID NO: 213.
XX
XX Human; secreted protein; gene therapy; vaccine; treatment; diagnosis;
KW GENSET; ss.
XX
OS Homo sapiens.
XX

PN WO200142451-A2.
XX
PD 14-JUN-2001.
XX
PF 07-DEC-2000; 2000WO-IB01938.
XX
PR 08-DEC-1999; 99US-0169629.
PR 06-MAR-2000; 2000US-0187470.
XX
XX (GEST) GENSET.
XX
PI Dumas Milne Edwards J, Bougueleret L, Jobert S;
XX
XX WPI; 2001-367870/38.
DR P-PSDB; AAG89334.
DR
XX
XX Full length GENSET human nucleic acids encoding potentially secreted
PT proteins, useful in gene therapy and vaccination against a variety of
PT diseases, and for diagnosis of those diseases -
XX
XX Claim 7; Page 766-767; 921pp; English.
XX
XX The invention relates to full length GENSET human nucleic acids encoding
CC potentially secreted proteins. The nucleic acids and the polypeptides
CC they encode may be used in the prevention, treatment and diagnosis of
CC diseases associated with inappropriate GENSET gene expression. For
CC example, they be used to treat disorders associated with decreased
CC GENSET gene expression by rectifying mutations or deletions in a
CC patient's genome that affect the activity of GENSET or by supplementing
CC the patient's own production of GENSET polypeptides. Conversely,
CC antisense nucleic acid molecules may be administered to down regulate
CC GENSET expression by binding with the cells' own genes and preventing
CC their expression. The sense and antisense nucleic acids may also be
CC used as DNA probes in diagnostic assays to detect and quantitate the
CC presence of similar nucleic acid sequences in samples, and hence to
CC determine which patients may be in need of restorative therapy.
CC The GENSET polypeptides may be used as antigens in the production of
CC antibodies and in assays to identify modulators (agonists and
CC antagonists) of GENSET polypeptide expression and activity. The
CC present sequence is a GENSET nucleic acid of the invention.
XX
SQ Sequence 1489 BP; 425 A; 323 C; 335 G; 406 T; 0 other;

Query Match 14.7%; Score 30.2; DB 22; Length 1489;
Best Local Similarity 50.3%; Pred. No. 5.3;
Matches 74; Conservative 0; Mismatches 73; Indels 0; Gaps 0;

QY 35 AGAACCCAGTTAAAGGTCCTGTGTCTACTAAGCCAGGTTCTTGTCTTATTATCTTGATTTCG 94
Db 145 AGAAAGGTTAAAGGCTGATGGTACCTAAAGCCTGGTACTTGAATTTGATCAAGATAAG 204
QY 95 TTGCGCTATGTTAAACCCACCTAACCGTTGTTGAAGGACACTGATTGTCAGGTATCAA 154
Db 205 CTGCCTTAAGTTCTCTTCATTACACAAATGATCCTAGATAATTGATAGATCCTGTGGTTC 264
QY 155 AAAGTGCTGTGAAGGTTCTCCTGCGGTAT 181
Db 265 AACTGGATTCTTAGATAGAAGCTGGAT 291

RESULT 52
ABN85329/C
ID ABN85329 standard; DNA; 36159 BP.
XX
AC ABN85329;
XX
DT 04-OCT-2002 (first entry)
XX
DE Human kinase gene.
XX
KW Human; kinase; enzyme; homeodomain-interacting protein kinase; testis;
KW brain medulloblastoma; infant brain; schizophrenia brain; retina; gene;
KW germinal center B cell; colon; liver; chromosome 1; gene therapy;
KW

DR P-PSDB; ABB83490.

XX Human kinase protein, related to homeodomain-interacting protein kinase

PT subfamily, useful as a model for developing human therapeutic targets

PT and serves as a target for human therapeutics

XX

PS Claim 22; Fig 3; 96pp; English.

XX

CC The present sequence is the gene for a human kinase, which is related to

CC the homeodomain-interacting protein kinase subfamily. The kinase and its

CC DNA sequence can be used as models for the development of human

CC therapeutic targets, aid in the identification of therapeutic proteins

CC and serve as targets for the development of human therapeutic agents that

CC modulate kinase activity in cells and tissues that express the kinase.

CC Experimental data indicates expression of the human kinase in testis,

CC brain medulloblastomas, infant brain, schizophrenic brain, retina,

CC germinal center B cells, colon and liver. The human kinase gene is

CC located on chromosome 1.

XX

SQ Sequence 36159 BP; 8425 A; 6177 C; 6348 G; 10756 T; 4453 other;

Query Match 14.7%; Score 30.2; DB 24; Length 36159;

Best Local Similarity 49.1%; Pred. No. 17;

Matches 80; Conservative 0; Mismatches 83; Indels 0; Gaps 0;

QY 31 CTCAGAACAAGTAAAGGTCCTGTCTACTAAGCCAGGTTCTTGTCTTATCTTGA 90

Db 12268 CTCAGATCCAGATCCAGATCCAGATTCAGATTCCTCATGTTACAGGAGTTAAATGGAAA 12209

QY 91 TTCGTTGCGTATGTTAAACCCACCTAACCGTTGTTTGAAGGACACTGATTGTCCAGGTA 150

Db 12208 ACAGATCTTTCGTFGAAGAACCATCTAACTGGCGGTGTCTCTCCACAGAATACCCAGTT 12149

QY 151 TCAAAAAGTGTCTGAAGGTTCTCTGCGGTATGCGTGTGTTTCGT 193

Db 12148 TCTCAAACTACTTTCAAAGTGATATGTGCTTATTGTTGTTTGT 12106

RESULT 53

AAK53717

ID AAK53717 standard; cDNA; 344 BP.

XX

AC AAK53717;

XX

DT 16-NOV-2001 (first entry)

XX

DE Murine replication associated protein encoding cDNA SEQ ID 282.

XX

KW Murine; liver; gene library; amino acid synthesis; binding protein;

KW cell metabolism; energy metabolism; fatty acid metabolism; synthesis;

KW phospholipid metabolism; purine; pyrimidine; nucleoside; nucleotide;

KW replication; transcription; translation; transport protein; ss.

XX

OS Mus musculus.

XX

PN DE20103510-U1.

XX

PD 07-JUN-2001.

XX

PF 28-FEB-2001; 2001DE-2003510.

XX

PR 02-DEC-1999; 99DE-1058160.

XX

PA (LION-) LION BIOSCIENCE AG.

XX

DR WPI; 2001-368570/39.

XX

PT Gene library containing sequences with specific 3'-ends and no polyA

PT tail, encoding proteins involved in a wide range of cellular processes

PT

XX 'Claim 15; Page 97; 251pp; German.

PS

XX

CC This invention describes a novel gene library (A) comprises a gene

CC sequence (or its part) encoding a protein involved in amino acid

CC synthesis, cellular/energy metabolism, metabolism of

CC fatty acids/phospholipids, synthesis or breakdown of

CC purines/pyrimidines/nucleosides/nucleotides, DNA

CC replication/transcription/translation, or is a transport/binding protein.

CC (A) are produced that correspond to the 3'-end of mRNA but without the

CC polyA tail. They can be prepared more efficiently and with less effort

CC than conventional libraries. AAK53436-AAK54275 represent fragments of the

CC gene library described in the method of the invention.

XX

SQ Sequence 344 BP; 73 A; 78 C; 70 G; 123 T; 0 other;

Query Match 14.5%; Score 29.8; DB 22; Length 344;

Best Local Similarity 56.7%; Pred. No. 4.2;

Matches 55; Conservative 0; Mismatches 42; Indels 0; Gaps 0;

QY 82 TTATCTTGATTCGTTGCGCTATGTTAAACCCACCTAACCGTTGTTTGAAGGACACTGATT 141

Db 230 TTAGCTGGTCTTTTCCCCCATGTTTATACAACCATATGTCCTTCGCTGGACCCCTGCTG 289

QY 142 GTCCAGGTATCAAAAAGTGTGTGAAGGTTCTCTGCGG 178

Db 290 ATCTGGATAGCAGAGATTTCGTCTCTGTGACCCGAGG 326

RESULT 54

AAK68952/c

ID AAT68952 standard; DNA; 5150 BP.

XX

AC AAT68952;

XX

DT 05-AUG-1997 (first entry)

XX

DE Blackcurrant fruit-specific RIB7 gene promoter.

XX

KW Blackcurrant; fruit-specific promoter; RIB7 gene; transgenic plant;

KW ss.

XX

OS Ribes nigrum strain Ben Alder.

XX

FH Key Location/Qualifiers

FT TATA_signal 3041..3044

FT /*tag= a

FT CDS 3156

FT /*tag= b

FT /codon_start= 3156

FT /note= "putative ATG translational start codon"

XX

PN WO9717452-A1.

XX

PD 15-MAY-1997.

XX

PF 04-NOV-1996; 96WO-EP04807.

XX

PR 03-NOV-1995; 95GB-0022558.

XX

PA (SMIK) SMITHKLINE BEECHAM PLC.

XX

PI Brennan RM, Taylor MA, Woodhead MR;

XX

DR WPI; 1997-281041/25.

XX

PT New isolated promoters from blackcurrant fruit - used for driving

PT fruit-specific expression of DNA sequences in transgenic

PT blackcurrant and other non-climacteric fruit

XX

PS Claim 3; Page 44-48; 66pp; English.

XX

CC A DNA sequence (AAT68952) comprises the promoter region of the

CC RIB7 gene. This gene was isolated from a blackcurrant genomic

CC library using a RIB7 cDNA clone (see also AAT68957) as probe. The

CC RIB7 gene is highly expressed in ripe blackcurrant fruit and

CC expressed at negligible levels in other tissues of the blackcurrant
CC plant. Its promoter region can be used to drive fruit-specific
CC expression of cloned downstream DNA sequences in transgenic
CC blackcurrant or other non-climacteric fruit as a means of
CC manipulating the ripening process.

XX
SQ Sequence 5150 BP; 1599 A; 1069 C; 899 G; 1583 T; 0 other;

Query Match 14.3%; Score 29.4; DB 18; Length 5150;
Best Local Similarity 55.3%; Pred. No. 16;
Matches 57; Conservative 0; Mismatches 46; Indels 0; Gaps 0;

QY 100 CTATGTTAAACCCACCTAACCGTTGTTTGAAGGACACTGATTGTCCAGGTATCAAAAAAGT 159
|||||
Db 2569 CTATTTAAACAAGACTTCTAAGTTGTTTCAGGGATTATTATTCATATTTTATCAATAAAA 2510
|||||

QY 160 GCTGTGAAGGTTCTCGGATGCGTGTGTTTTCGTTCCACAATA 202
|||||
Db 2509 ACTGAGAACTTTTCTCCGTTGTTCTTCTTCTTACGAGAATA 2467
|||||

RESULT 55
ABQ89443
ID ABQ89443 standard; cDNA; 756 BP.
XX
AC ABQ89443;
XX
DT 27-SEP-2002 (first entry)
XX
DE Human prostate expressed polynucleotide SEQ ID NO 699.
XX
KW Human; prostate; cytostatic; tumour; cancer; vaccine; gene therapy;
KW gene; ss.
XX
OS Homo sapiens.
XX
PN WO200255700-A2.
XX
PD 18-JUL-2002.
XX
PF 07-DEC-2001; 2001WO-US47349.
XX
PR 07-DEC-2000; 2000US-254648P.
PR 13-MAR-2001; 2001US-275688P.
XX
PA (CHIR) CHIRON CORP.
PA (HYSE-) HYSEQ INC.
XX
PI Escobedo J, Garcia PD, Kassam A, Lamson G, Drmanac R;
PI Crkvenjakov R, Dickson M, Drmanac S, Labat I, Leshkowitz D;
PI Kita D, Garcia V, Jones WL, Stache-Grain B, Scott EM;
XX
DR WPI; 2002-557824/59.
XX
PT New genes and gene products isolated from human prostate, useful for
PT treating or diagnosing tumor or cancer (e.g. prostate cancer or breast
PT cancer), or as vaccines for treating or preventing these diseases -
XX
PS Claim 1; SEQ ID NO 699; 186pp + Sequence Listing; English.
XX
CC The invention relates to an isolated polynucleotide comprising any of
CC 1477 sequences or its fragment, degenerate variant, antisense or
CC complement. The polynucleotides and gene products are useful for treating
CC or diagnosing tumour or cancer (e.g. prostate cancer, breast cancer,
CC lung cancer or medullary carcinoma) in a subject (e.g. cattle, dogs,
CC cats, rabbits, horse or human). The polynucleotides and polypeptides are
CC also useful as vaccines for treating or preventing these diseases. The
CC polynucleotides are useful for gene therapy. The present sequence is that
CC of one of a group of polynucleotides (ABQ88745-ABQ90015) disclosed
CC electronically as sequences of the invention. However only 1271
CC polynucleotide sequences are given, whereas 1477 polynucleotides and 91
CC proteins are claimed.
CC Note: The sequence data for this patent did not form part of the printed

CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequence.

XX
SQ Sequence 756 BP; 236 A; 138 C; 158 G; 214 T; 10 other;

Query Match 14.2%; Score 29.2; DB 24; Length 756;
Best Local Similarity 49.4%; Pred. No. 9;
Matches 76; Conservative 0; Mismatches 78; Indels 0; Gaps 0;

QY 38 ACCAGTTAAAGGTCCTGTCTACTAAGCCAGGTTCTTCTCTATTATCTTGATTCGTTG 97
|||||
Db 291 ACTATTAAATCATCTGGCTTACTAAACAAACCTTTTCTGCACTGACTTCTTTCTGTG 350
|||||

QY 98 CGCTATGTTAAACCCACCTAACCGTTGTTTGAAGGACACTGATTGTCCAGGTATCAAAAA 157
|||||
Db 351 TGTGATTCAAATTAATTTTAATGTTTACTAAAAACAAAGTTATCTACCATTTATTATAAA 410
|||||

QY 158 GTGCTGTGAAGGTTCTCTGCGGTATGCGTTGTTTC 191
|||||
Db 411 ACTCAATTGTAGGCCAGGCGTGGTGGCTTACGTC 444
|||||

RESULT 56
AAS00193/c
ID AAS00193 standard; DNA; 549 BP.
XX
AC AAS00193;
XX
DT 04-JUL-2001 (first entry)
XX
DE S. aureus DNA encoding transcription antitermination factor, NusG.
XX
KW Transcription antitermination factor; NusG; immunogen; vaccine;
KW antibody; wound infection; cellulitis; burn infection; eyelid infection;
KW food poisoning; joint infection; neonatal conjunctivitis; osteomyelitis;
KW skin infection; scalded skin syndrome; toxic epidermal necrosis;
KW Ritter's disease; Lyell's disease; toxic shock syndrome; endocarditis;
KW ds.
XX
OS Staphylococcus aureus.
XX
FH Key Location/Qualifiers
FT CDS 1..549
FT /*tag= a
FT /product= "NusG"
FT /partial
FT /note= "No stop codon"

PN WO200116292-A2.
XX
PD 08-MAR-2001.
XX
PF 31-AUG-2000; 2000WO-US23773.
XX
PR 01-SEP-1999; 99US-0151933.
XX
PA (HUMA-) HUMAN GENOME SCI INC.
XX
PI Choi GH;
XX
DR WPI; 2001-183259/18.
DR P-PSDB; AAU00832.
XX
PT New isolated nucleic acid for use in diagnosing Staphylococcus
PT infections and in vaccines for eliciting immune responses to the
PT infections -
XX
PS Claim 1; Page 15; 225pp; English.
XX
CC The sequence encodes S. aureus NusG (transcription antitermination
CC factor). The polynucleotides of the invention are used to detect
CC Staphylococcus nucleic acids in a biological sample from an animal for
CC diagnosing Staphylococcus infections. The polypeptides of the invention

PR 14-JUL-2000; 2000US-0218290.
PR 26-JUL-2000; 2000US-0220963.
PR 26-JUL-2000; 2000US-0220964.
PR 14-AUG-2000; 2000US-0224518.
PR 14-AUG-2000; 2000US-0224519.
PR 14-AUG-2000; 2000US-0225213.
PR 14-AUG-2000; 2000US-0225214.
PR 14-AUG-2000; 2000US-0225266.
PR 14-AUG-2000; 2000US-0225267.
PR 14-AUG-2000; 2000US-0225268.
PR 14-AUG-2000; 2000US-0225270.
PR 14-AUG-2000; 2000US-0225447.
PR 14-AUG-2000; 2000US-0225757.
PR 14-AUG-2000; 2000US-0225758.
PR 14-AUG-2000; 2000US-0225759.
PR 18-AUG-2000; 2000US-0226279.
PR 22-AUG-2000; 2000US-0226681.
PR 22-AUG-2000; 2000US-0226868.
PR 22-AUG-2000; 2000US-0227182.
PR 23-AUG-2000; 2000US-0227009.
PR 30-AUG-2000; 2000US-0228924.
PR 01-SEP-2000; 2000US-0229287.
PR 01-SEP-2000; 2000US-0229343.
PR 01-SEP-2000; 2000US-0229344.
PR 01-SEP-2000; 2000US-0229345.
PR 05-SEP-2000; 2000US-0229509.
PR 05-SEP-2000; 2000US-0229513.
PR 06-SEP-2000; 2000US-0230437.
PR 06-SEP-2000; 2000US-0230438.
PR 08-SEP-2000; 2000US-0231242.
PR 08-SEP-2000; 2000US-0231243.
PR 08-SEP-2000; 2000US-0231244.
PR 08-SEP-2000; 2000US-0231413.
PR 08-SEP-2000; 2000US-0231414.
PR 08-SEP-2000; 2000US-0232080.
PR 08-SEP-2000; 2000US-0232081.
PR 12-SEP-2000; 2000US-0231968.
PR 14-SEP-2000; 2000US-0232397.
PR 14-SEP-2000; 2000US-0232398.
PR 14-SEP-2000; 2000US-0232399.
PR 14-SEP-2000; 2000US-0232400.
PR 14-SEP-2000; 2000US-0232401.
PR 14-SEP-2000; 2000US-0233063.
PR 14-SEP-2000; 2000US-0233064.
PR 14-SEP-2000; 2000US-0233065.
PR 21-SEP-2000; 2000US-0234223.
PR 21-SEP-2000; 2000US-0234274.
PR 25-SEP-2000; 2000US-0234997.
PR 25-SEP-2000; 2000US-0234998.
PR 26-SEP-2000; 2000US-0235484.
PR 27-SEP-2000; 2000US-0235834.
PR 27-SEP-2000; 2000US-0235836.
PR 29-SEP-2000; 2000US-0236327.
PR 29-SEP-2000; 2000US-0236367.
PR 29-SEP-2000; 2000US-0236368.
PR 29-SEP-2000; 2000US-0236369.
PR 29-SEP-2000; 2000US-0236370.
PR 02-OCT-2000; 2000US-0236802.
PR 02-OCT-2000; 2000US-0237037.
PR 02-OCT-2000; 2000US-0237038.
PR 02-OCT-2000; 2000US-0237039.
PR 02-OCT-2000; 2000US-0237040.
PR 13-OCT-2000; 2000US-0239935.
PR 13-OCT-2000; 2000US-0239937.
PR 20-OCT-2000; 2000US-0240960.
PR 20-OCT-2000; 2000US-0241221.
PR 20-OCT-2000; 2000US-0241785.
PR 20-OCT-2000; 2000US-0241786.
PR 20-OCT-2000; 2000US-0241787.
PR 20-OCT-2000; 2000US-0241808.
PR 20-OCT-2000; 2000US-0241809.
PR 20-OCT-2000; 2000US-0241826.
PR 01-NOV-2000; 2000US-02444617.

PR 08-NOV-2000; 2000US-0246474.
PR 08-NOV-2000; 2000US-0246475.
PR 08-NOV-2000; 2000US-0246476.
PR 08-NOV-2000; 2000US-0246477.
PR 08-NOV-2000; 2000US-0246478.
PR 08-NOV-2000; 2000US-0246523.
PR 08-NOV-2000; 2000US-0246524.
PR 08-NOV-2000; 2000US-0246525.
PR 08-NOV-2000; 2000US-0246526.
PR 08-NOV-2000; 2000US-0246527.
PR 08-NOV-2000; 2000US-0246528.
PR 08-NOV-2000; 2000US-0246532.
PR 08-NOV-2000; 2000US-0246609.
PR 08-NOV-2000; 2000US-0246610.
PR 08-NOV-2000; 2000US-0246611.
PR 08-NOV-2000; 2000US-0246613.
PR 17-NOV-2000; 2000US-0249207.
PR 17-NOV-2000; 2000US-0249208.
PR 17-NOV-2000; 2000US-0249209.
PR 17-NOV-2000; 2000US-0249210.
PR 17-NOV-2000; 2000US-0249211.
PR 17-NOV-2000; 2000US-0249212.
PR 17-NOV-2000; 2000US-0249213.
PR 17-NOV-2000; 2000US-0249214.
PR 17-NOV-2000; 2000US-0249215.
PR 17-NOV-2000; 2000US-0249216.
PR 17-NOV-2000; 2000US-0249217.
PR 17-NOV-2000; 2000US-0249218.
PR 17-NOV-2000; 2000US-0249244.
PR 17-NOV-2000; 2000US-0249245.
PR 17-NOV-2000; 2000US-0249264.
PR 17-NOV-2000; 2000US-0249265.
PR 17-NOV-2000; 2000US-0249297.
PR 17-NOV-2000; 2000US-0249299.
PR 17-NOV-2000; 2000US-0249300.
PR 01-DEC-2000; 2000US-0250160.
PR 01-DEC-2000; 2000US-0250391.
PR 05-DEC-2000; 2000US-0251030.
PR 05-DEC-2000; 2000US-0251988.
PR 05-DEC-2000; 2000US-0256719.
PR 06-DEC-2000; 2000US-0251479.
PR 08-DEC-2000; 2000US-0251856.
PR 08-DEC-2000; 2000US-0251868.
PR 08-DEC-2000; 2000US-0251869.
PR 08-DEC-2000; 2000US-0251989.
PR 08-DEC-2000; 2000US-0251990.
PR 11-DEC-2000; 2000US-0254097.
PR 05-JAN-2001; 2001US-0259678.
XX
PA
XX
PI
XX
DR
XX
PT
PT
XX
PS
XX
CC
CC
CC
CC
CC
CC
CC
CC
CC
CC
CC

(HUMA-) HUMAN GENOME SCI INC.

Rosen CA, Barash SC, Ruben SM;

WPI; 2001-483426/52.

Nucleic acids encoding human immune/hematopoietic antigen polypeptides, useful for preventing, diagnosing and/or treating cancers and metastasis -

Disclosure; SEQ ID NO 23387; 3071pp + Sequence Listing; English.

AAK54951 to AAK64702 encode the human immune/haematopoietic antigen (I) amino acid sequences given in AAM82170 to AAM91921. (I) have cytostatic activity, and can be used in gene therapy and vaccine production. (I) proteins and polynucleotides may be used in the prevention, diagnosis and treatment of diseases associated with inappropriate (I) expression. For example, they may be used to treat disorders associated with decreased expression by rectifying mutations or deletions in a patient's genome that affect the activity of (I) by expressing inactive proteins or to supplement the patients own production of (I). Additionally, (I) polynucleotides may be used to produce the secreted (I), by inserting the nucleic acids into a host cell and culturing the cell to express the protein. (I) proteins and polynucleotides may be used to prevent,

Qy 137 TGATTGTCCAGGTATCAAAAAAGTGTCTGTGAAGGTTCTCTGGGG 178

Db 353 TGATTTTCTGGCATCATGCATGACTGAGAAGGTCAGTGTTGG 394

RESULT 65
ABA23689
ID ABA23689 standard; DNA: 464 BP.

RESULT 66
AAK02208
ID AAK02208 standard; DNA; 464 BP.
XX
AC AAK02208;
XX
DT 05-NOV-2001 (first entry)
XX
DE Human brain expressed single exon probe SEQ ID NO: 2199.
XX
Human; brain expressed exon; gene expression analysis; probe;
KW microarray; Alzheimer's disease; multiple sclerosis; schizophrenia;
KW epilepsy; cancer; ss.

[illegible]

DE Human bone marrow expressed single exon probe SEQ ID NO: 2212.
XX Human; bone marrow expressed exon; gene expression analysis; probe;
KW microarray; cancer; leukaemia; lymphoma; myeloma; ss.
XX Homo sapiens.
OS WO200157276-A2.
PN 09-AUG-2001.
XX 30-JAN-2001; 2001WO-US00668.
PF 04-FEB-2000; 2000US-0180312.
XX 26-MAY-2000; 2000US-0207456.
PR 30-JUN-2000; 2000US-0608408.
PR 03-AUG-2000; 2000US-0632366.
PR 21-SEP-2000; 2000US-0234687.
PR 27-SEP-2000; 2000US-0236359.
PR 04-OCT-2000; 2000GB-0024263.
XX (MOLE-) MOLECULAR DYNAMICS INC.
PA Penn SG, Hanzel DK, Chen W, Rank DR;
PI WPI; 2001-488900/53.
XX Human genome-derived single exon nucleic acid probes useful for
PT analyzing gene expression in human bone marrow -
XX Example 4; SEQ ID NO: 2212; 658pp + Sequence Listing; English.
PS The present invention provides a number of single exon nucleic acid
XX probes which are derived from genomic sequences expressed in the human
CC bone marrow. They can be used to measure gene expression in bone marrow
CC samples, which may enable the improved diagnosis and treatment of cancers
CC such as lymphoma, leukaemia and myeloma. The present sequence is one of
XX the probes of the invention.
SQ Sequence 464 BP; 112 A; 126 C; 88 G; 138 T; 0 other;
Query Match 13.8%; Score 28.4; DB 22; Length 464;
Best Local Similarity 54.9%; Pred. No. 14;
Matches 56; Conservative 0; Mismatches 46; Indels 0; Gaps 0;
QY 77 TCCTATTATCTTGGCTATGTTAAACCCACCTAACCGTTGTTGAAGGACAC 136
|||||
Db 293 TCCTCTTCCACAGATTATTGCCTGCTCCTCCAAAGCCCTTCACTGTAGATGAAGCCTG 352
QY 137 TGATTGTCCAGGTATCAAAAAGTGTGTGAAGGTTCTCTGCGG 178
|||||
Db 353 TGATTTTCTGGCATCATGCTGACTGAGAAGGTCAGTGTGG 394
RESULT 68
AAI12235
ID AAI12235 standard; DNA; 464 BP.
XX AAI12235;
AC 12-OCT-2001 (first entry)
DT Probe #2168 for gene expression analysis in human cervical cell sample.
XX Probe; human; microarray; gene expression; cervical epithelial cell;
DE cervical cancer; ss.
KW Homo sapiens.
XX WO200157278-A2.
PN 09-AUG-2001.
XX

PF 30-JAN-2001; 2001WO-US00670.
XX 04-FEB-2000; 2000US-0180312.
PR 26-MAY-2000; 2000US-0207456.
PR 30-JUN-2000; 2000US-0608408.
PR 03-AUG-2000; 2000US-0632366.
PR 21-SEP-2000; 2000US-0234687.
PR 27-SEP-2000; 2000US-0236359.
PR 04-OCT-2000; 2000GB-0024263.
XX (MOLE-) MOLECULAR DYNAMICS INC.
PA Penn SG, Hanzel DK, Chen W, Rank DR;
PI WPI; 2001-488901/53.
XX Human genome-derived single exon nucleic acid probes useful for
PT analyzing gene expression in human cervical epithelial cells -
XX Claim 25; SEQ ID No 2168; 487pp; English.
PS The present invention relates to human single exon nucleic acid probes
XX (SENPs). The present sequence is one such probe. The SENPs are derived
CC from human HeLa cells. The SENPs can be used to produce a single exon
CC microarray, which can be used for measuring human gene expression in a
CC sample derived from human cervical epithelial cells. By measuring gene
CC expression, the probes are therefore useful in grading and/or staging
CC of diseases of the cervix, notably cervical cancer.
CC Note: The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 464 BP; 112 A; 126 C; 88 G; 138 T; 0 other;
Query Match 13.8%; Score 28.4; DB 22; Length 464;
Best Local Similarity 54.9%; Pred. No. 14;
Matches 56; Conservative 0; Mismatches 46; Indels 0; Gaps 0;
QY 77 TCCTATTATCTTGGCTATGTTAAACCCACCTAACCGTTGTTGAAGGACAC 136
|||||
Db 293 TCCTCTTCCACAGATTATTGCCTGCTCCTCCAAAGCCCTTCACTGTAGATGAAGCCTG 352
QY 137 TGATTGTCCAGGTATCAAAAAGTGTGTGAAGGTTCTCTGCGG 178
|||||
Db 353 TGATTTTCTGGCATCATGCTGACTGAGAAGGTCAGTGTGG 394
RESULT 69
AAI33589
ID AAI33589 standard; DNA; 464 BP.
XX AAI33589;
AC 17-OCT-2001 (first entry)
DT Probe #2275 used to measure gene expression in human placenta sample.
XX Probe; microarray; human; placenta; antenatal diagnosis;
DE genetic disorder; ss.
KW Homo sapiens.
XX WO200157272-A2.
PN 09-AUG-2001.
XX 30-JAN-2001; 2001WO-US00663.
PF 04-FEB-2000; 2000US-0180312.
XX 26-MAY-2000; 2000US-0207456.
PR 30-JUN-2000; 2000US-0608408.
PR 03-AUG-2000; 2000US-0632366.
PR 21-SEP-2000; 2000US-0234687.

PR 27-SEP-2000; 2000US-0236359.
PR 04-OCT-2000; 2000GB-0024263.
XX
PA (MOLE-) MOLECULAR DYNAMICS INC.
XX
PI Penn SG, Hanzel DK, Chen W, Rank DR;
XX WPI; 2001-488997/53.
DR
XX Human genome-derived single exon nucleic acid probes useful for
XX analyzing gene expression in human placenta -
PT
XX Claim 25; SEQ ID No 2275; 654pp; English.
PS
XX The present invention relates to single exon nucleic acid probes (SENPs).
CC The present sequence is one such probe. The probes are useful for
CC producing a microarray for predicting, measuring and displaying gene
CC expression in samples derived from human placenta. The probes are useful
CC for antenatal diagnosis of human genetic disorders.
XX
SQ Sequence 464 BP; 112 A; 126 C; 88 G; 138 T; 0 other;
Query Match 13.8%; Score 28.4; DB 22; Length 464;
Best Local Similarity 54.9%; Pred. No. 14;
Matches 56; Conservative 0; Mismatches 46; Indels 0; Gaps 0;
QY 77 TCCTATTATCTTGATTCGTTGCGCTATGTTAAACCCACCTAACCGTTGTTTGAAGGACAC 136
Db 293 TCCTCTTTCCACAGATTATTGCTGCTCTCCTCAAGCCCTTCACTGTAGATGAAGCCTG 352
QY 137 TGATTGTCCAGGTATCAAAAAGTGTGTGAAGGTTCTCTGCGG 178
Db 353 TGATTTCCTGGCATCATGCTGACTGAGAGGTCAGTGTGG 394
RESULT 70
AAI02148
ID AAI02148 standard; DNA; 464 BP.
XX
AC AAI02148;
XX
DT 09-OCT-2001 (first entry)
XX
DE Probe #2139 used to measure gene expression in human breast sample.
XX
KW Probe; human; breast disease; breast cancer; development disorder; ss;
KW inflammatory disease; proliferative breast disease; non-carcinoma tumour.
XX
OS Homo sapiens.
XX
PN WO200157270-A2.
XX
PD 09-AUG-2001.
XX
PF 29-JAN-2001; 2001WO-US00661.
XX
PR 04-FEB-2000; 2000US-0180312.
PR 26-MAY-2000; 2000US-0207456.
PR 30-JUN-2000; 2000US-0608408.
PR 03-AUG-2000; 2000US-0632366.
PR 21-SEP-2000; 2000US-0234687.
PR 27-SEP-2000; 2000US-0236359.
PR 04-OCT-2000; 2000GB-0024263.
XX
PA (MOLE-) MOLECULAR DYNAMICS INC.
XX
PI Penn SG, Hanzel DK, Chen W, Rank DR;
XX WPI; 2001-476286/51.
XX
PT Novel single exon nucleic acid probe used to measuring gene expression
PT 'in a human breast -
XX

PS Claim 25; SEQ ID No 2139; 322pp; English.
XX
CC The present invention relates to novel single exon nucleic acid probes.
CC The present sequence is one such probe. The probes are useful for
CC measuring human gene expression in a human breast sample, where the probe
CC hybridises at high stringency to a nucleic acid expressed in the human
CC breast. The probes are useful for predicting, diagnosing, grading,
CC staging, monitoring and prognosing diseases of the human breast,
CC particularly those diseases with polygenic aetiology. The diseases
CC include: breast cancer, disorders of development, inflammatory diseases
CC of the breast, fibrocystic changes, proliferative breast disease and
CC non-carcinoma tumours.
CC Note: The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 464 BP; 112 A; 126 C; 88 G; 138 T; 0 other;
Query Match 13.8%; Score 28.4; DB 22; Length 464;
Best Local Similarity 54.9%; Pred. No. 14;
Matches 56; Conservative 0; Mismatches 46; Indels 0; Gaps 0;
QY 77 TCCTATTATCTTGATTCGTTGCGCTATGTTAAACCCACCTAACCGTTGTTTGAAGGACAC 136
Db 293 TCCTCTTTCCACAGATTATTGCTGCTCTCCTCAAGCCCTTCACTGTAGATGAAGCCTG 352
QY 137 TGATTGTCCAGGTATCAAAAAGTGTGTGAAGGTTCTCTGCGG 178
Db 353 TGATTTCCTGGCATCATGCTGACTGAGAGGTCAGTGTGG 394
RESULT 71
ABS02116
ID ABS02116 standard; DNA; 464 BP.
XX
AC ABS02116;
XX
DT 19-AUG-2002 (first entry)
XX
DE Human genome-derived single exon probe from lung SEQ ID No 2107.
XX
KW Human; ds; single exon probe; asthma; lung cancer; COPD; ILD;
KW chronic obstructive pulmonary disease; interstitial lung disease;
KW familial idiopathic pulmonary fibrosis; neurofibromatosis;
KW tuberous sclerosis; Gaucher's disease; Niemann-Pick disease;
KW Hermansky-Pudlak syndrome; sarcoidosis; pulmonary haemosiderosis;
KW pulmonary histiocytosis; lymphangioleiomyomatosis; Karagener syndrome;
KW pulmonary alveolar proteinosis; fibrocystic pulmonary dysplasia;
KW primary ciliary dyskinesia; pulmonary hypertension;
KW hyaline membrane disease.
XX
OS Homo sapiens.
XX
PN WO200186003-A2.
XX
PD 15-NOV-2001.
XX
PF 30-JAN-2001; 2001WO-US00665.
XX
PR 04-FEB-2000; 2000US-180312P.
PR 26-MAY-2000; 2000US-207456P.
PR 30-JUN-2000; 2000US-0608408.
PR 03-AUG-2000; 2000US-0632366.
PR 21-SEP-2000; 2000US-234687P.
PR 27-SEP-2000; 2000US-236359P.
PR 04-OCT-2000; 2000GB-0024263.
XX
PA (MOLE-) MOLECULAR DYNAMICS INC.
XX
PI Penn SG, Hanzel DK, Chen W, Rank DR;
XX WPI; 2002-114183/15.
XX

XX Human; secreted protein; ss; antiinflammatory; immunosuppressive;
KW nootropic; neuroprotective; antiarthritic; antimicrobial; vulnerary;
KW cyostatic; antidiabetic; virucide; antiinfertility; anticonvulsant;
KW vasotropic; antiparkinsonian; immunostimulant; dermatological;
KW antirheumatic; antitumor; antiulcer; osteopathic; tranquiliser;
KW cerebroprotective; cytokine; cell proliferation; cell differentiation;
KW immune deficiency; severe combined immunodeficiency; SCID; tumour;
KW autoimmune disorder; multiple sclerosis; rheumatoid arthritis;
KW graft-versus-host disease; myeloid deficiency; wound healing; ulcer;
KW periodontal disease; osteoporosis; osteoarthritis; Alzheimer's disease;
KW Parkinson's disease; Huntington's disease; infection; cardiac disease;
KW stroke; sepsis; inflammatory bowel disease; contraceptive; immunogen;
KW food supplement.
XX
OS Homo sapiens.
XX
PN WO200175068-A2.
XX
PD 11-OCT-2001.
XX
PF 22-MAR-2001; 2001WO-US09369.
XX
PR 30-MAR-2000; 2000US-0539330.
PR 04-DEC-2000; 2000US-0729674.
XX
PA (GEMY) GENETICS INST INC.
XX
PI Jacobs K, McCoy JM, Lavallie E, Collins-racie LA, Evans C;
PI Treacy M, Agostino MJ, Steininger RJ, Spaulding V, Wong GG;
PI Clark H, Fechtel K, Merberg D;
XX
DR WPI; 2001-639363/73.
DR P-PSDB; AAU38998.
XX
PT Secreted human proteins, useful as vaccine for treating various
PT diseases such as autoimmune disorders (e.g. multiple sclerosis), and
PT nervous system disorders (e.g. stroke) -
XX
PS Claim 13; Page 471; 619pp; English.
XX
CC The invention relates to novel human secreted proteins, the nucleic
CC acids encoding them. The protein may exhibit cytokine, cell proliferation
CC or cell differentiation activity or may induce production of other
CC cytokines in certain cell populations and may exhibit immune stimulating
CC or immune suppressing activity, which is useful for the treatment of
CC various immune deficiencies and disorders e.g. severe combined
CC immunodeficiency (SCID), autoimmune disorders e.g. multiple sclerosis,
CC systemic lupus erythematosus, rheumatoid arthritis, autoimmune pulmonary
CC inflammation. The proteins are also useful in the treatment of diseases
CC and disorders including tissue, skin and organ transplantation and in
CC graft-versus-host diseases (GVHD), in the induction of tumour immunity,
CC myeloid or lymphoid cell deficiencies, wound healing and tissue repair,
CC in the treatment of burns, incisions and ulcers; as well as in treatment
CC of periodontal disease, osteoporosis or osteoarthritis, mediated by
CC inflammatory processes, diseases of the peripheral nervous system,
CC Alzheimer's, Parkinson's disease, Huntington's disease,
CC amyotrophic lateral sclerosis, and Shy-Drager syndrome, infections,
CC infarction of cardiac and central nervous system vessel e.g. stroke,
CC sepsis, inflammatory bowel disease, ulcers, bone regeneration. The
CC protein, having activin- or inhibin-related activities is useful as a
CC contraceptive based on the ability of inhibins to decrease fertility in
CC female mammals and decrease spermatogenesis in male mammals. The
CC proteins and nucleic acids are also useful as food supplements. The
CC present sequence encodes a secreted protein of the invention.
XX
SQ Sequence 1656 BP; 473 A; 389 C; 340 G; 454 T; 0 other;

Query Match 13.8%; Score 28.4; DB 22; Length 1656;
Best Local Similarity 62.9%; Pred. No. 22;
Matches 44; Conservative 0; Mismatches 26; Indels 0; Gaps 0;
QY 32 TCAAGAACAGGTAAAGGTCCTGTGTCTACTAAGCAGGTTCTTGTCTATTATCTTGAT 91

Db 1023 TCAAGCTACAGGTAAAGGCTCTACCACTTAAGGTATTGACTTATCCTATGAATTGAT 964
Qy 92 TCGTTGCGCT 101
Db 963 TCCTTCTCT 954
Search completed: February 15, 2003, 22:16:54
Job time : 811 secs

